TITLE: IDENTIFYING THE INVISIBLE IMPACT OF SCHOLARLY PUBLICATIONS: A MULTI-DISCIPLINARY ANALYSIS USING ALTMETRICS

A thesis submitted in partial fulfilment of the

requirements of the University of Wolverhampton

for the degree of Doctor of Philosophy

November 2014

This work or any part thereof has not previously been presented in any form to the University or to any other body whether for the purposes of assessment, publication or for any other purpose (unless otherwise indicated). Save for any express acknowledgments, references and/or bibliographies cited in the work, I confirm that the intellectual content of the work is the result of my own efforts and of no other person.

Signature ………………………………………..

Date …………………………………………….....

Abstract:

The field of 'altmetrics' is concerned with alternative metrics for the impact of research publications using social web data. Empirical studies are needed, however, to assess the validity of altmetrics from different perspectives. This thesis partly fills this gap by exploring the suitability and reliability of two altmetrics resources: Mendeley, a social reference manager website, and Faculty of F1000 (F1000), a post- publishing peer review platform. This thesis explores the correlations between the new metrics and citations at the level of articles for several disciplines and investigates the contexts in which the new metrics can be useful for research evaluation across different fields.

Low and medium correlations were found between Mendeley readership counts and citations for Social Sciences, Humanities, Medicine, Physics, Chemistry and Engineering articles from the Web of Science (WoS), suggesting that Mendeley data may reflect different aspects of research impact. A comparison between information flows based on Mendeley bookmarking data and cross-disciplinary citation analysis for social sciences and humanities disciplines revealed substantial similarities and some differences. This suggests that Mendeley readership data could be used to help identify knowledge transfer between scientific disciplines, especially for people that read but do not author articles, as well as providing evidence of impact at an earlier stage than is possible with citation counts.

The majority of Mendeley readers for Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry papers were PhD students

and postdocs. The highest correlations between citations and Mendeley readership counts were for types of Mendeley users that often authored academic papers, suggesting that academics bookmark papers in Mendeley for reasons related to scientific publishing.

In order to identify the extent to which Mendeley bookmarking counts reflect readership and to establish the motivations for bookmarking scientific papers in Mendeley, a large-scale survey found that 83% of Mendeley users read more than half of the papers in their personal libraries. The main reasons for bookmarking papers were citing in future publications, using in professional activities, citing in a thesis, and using in teaching and assignments. Thus, Mendeley bookmarking counts can potentially indicate the readership impact of research papers that have educational value for non-author users inside academia or the impact of research papers on practice for readers outside academia.

This thesis also examines the relationship between article types (i.e., “New Finding”, “Confirmation”, “Clinical Trial”, “Technical Advance”, “Changes to Clinical Practice”, “Review”, “Refutation”, “Novel Drug Target”), citation counts and F1000 article factors (FFa). In seven out of nine cases, there were no significant differences between article types in terms of rankings based on citation counts and the F1000 Article Factor (FFa) scores. Nevertheless, citation counts and FFa scores were significantly different for articles tagged: “New finding” or “Changes to Clinical Practice”. This means that F1000 could be used in research evaluation exercises when the importance of practical findings needs to be recognised. Furthermore, since the majority of the studied articles were reviewed in their year of publication, F1000 could also be useful for quick evaluations.

Table of Contents

List of Tables

List of Figures

List of Appendices

Publications from this Thesis

Journal articles

- Mohammadi, E., Thelwall, M., Kousha, K. (in preparation). Motivations for bookmarking in Mendeley.
- Mohammadi, E., Thelwall, M., Haustein, S., & Larivière, V. (in press). Who reads research articles? An altmetrics analysis of Mendeley user categories. Journal of the Association for Information Science and Technology.
- Mohammadi, E. & Thelwall, M. (in press). Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. Journal of the Association for Information Science and Technology.
- Mohammadi, E., & Thelwall, M. (2013). Assessing non-standard article impact using F1000 labels. Scientometrics. 97(2), 383-395.

Conference proceedings

- Mohammadi, E. & Thelwall, M. (2013). Assessing the Mendeley readership of social sciences and humanities research. 14th International Society of Scientometrics and Informetrics Conference, Vienna, Austria. 16-19th July 2013.

Acknowledgements

I dedicate this thesis to the memory of my mother and grandmother.

Special thanks to my lovely wife, Azar, for not only coping with the challenges of student life but also for finding solutions to have a smooth life during writing this thesis.

Thanks to Professor Mike Thelwall for being more than a supervisor; he taught me many lessons beyond this thesis.

Thanks to my second supervisor, Dr. Kayvan Kousha, for his fruitful and critical thoughts on this project.

Thanks to my dad for inspring me by his daily phone calls during the PhD. I would also like to thank my brother, Alborz, and my old friened, Mehdi, for their support.

I would also like to thank all of my colleagues for their very useful and inspirational feedback on my project, especially Dr. Jonathan Levitt, Dr. Stefanie Haustein and Dr.Vincent Larivière.

CHAPTER 1: INTRODUCTION

1.1 Background

Measuring the impact of scholarly publications has been a concern for academia, science policy makers and funding bodies for many years. In response, the field of scientometrics emerged in the 1960s to study scholarly communication with quantitative methods from different perspectives (Hood & Wilson, 2001). Consequently, evaluative bibliometrics developed with a similar focus on assessing scientific activities (Narin, 1976). At the same time, the Institute for Scientific Information (ISI) enabled a new direction for research evaluation from scientometric perspectives by establishing the Science Citation Index. As a result, the Journal Impact Factor (JIF) based on citation data became a relatively well-known indicator for measuring the impact of scientific publications. Because the JIF cannot reflect all dimensions of journals evaluation (Braun, 2012), it has been used inappropriately for assessing the research performance of researchers, institutions and research groups (Opthof, 1997; Boell & Wilson, 2010). Scientometricians have also focused on different levels of analysis for research evaluation. Several new citation-based metrics have been developed for evaluating research output at the level of research fields, such as the crown indicator (Waltman, van Eck, van Leeuwen, Visser, & van Raan, 2011), and authors (Hirsch, 2005; Egghe, 2006) during recent decades. This shows the progress of scientometrics in research assessment (Kamalski, Plume, & Mayur, 2014). Thus, the different indicators provide broader possibilities for scientometricians to evaluate research publications more effectively.

Although citation analysis methods have developed significantly, citation analysis has limitations for research assessment (MacRoberts & MacRoberts, 1989;

MacRoberts & MacRoberts, 1996). This is not a new challenge to the bibliometric community and researchers can try to partly fill this gap with new types datasets in addition to citations (limitations of citation are discussed in Chapter 2 Section 2.1.2). Researchers have used the web as a data source to examine scholarly communication in both formal and informal contexts (Thelwall, 2008). Related studies using webometrics methods are discussed in detail in Chapter 2 Sections 2.2.1.

The impact of research papers is multidimensional in nature (Martin, 1996; Sutherland, Goulson, Potts, & Dicks, 2011) and it is not reasonable to rely on a single metric to capture all of the impact of research publications (Moed et al., 2012). It is clear that the impacts of research can go beyond knowledge advancement within science, and, therefore, the influence of research publications in social, economic, cultural and environmental contexts needs to be identified (Bornmann, 2012; Thelwall, 2012). With the growth of informal scholarly communication on the web, it would be appropriate to partly assess the next generation of scholars in ways other than through traditional publications (Cronin, 2014). The Higher Education Funding Council for England (HEFCE), in the new Research Excellence Framework (REF), considers all types of research impacts such as "social, economic and cultural benefits and impacts beyond academia" (HEFCE, 2011, p.4). The National Science Foundation (NSF) already asks principal investigators to submit their research products for new grant applications (from 14 January 2013). This means that the NSF will not evaluate researchers based on their publications only, but also that non-conventional products of research are considered (NSF, 2013) and "it will alter how scientists assess research impact" (Piwowar, 2013, p.159). The ACUMEN (Academic Careers Understood through Measurement and Norms) Portfolio also recommended webometric and altmetric indicators in addition to bibliometric measures for the research evaluation of academics in the EU (http://research-acumen.eu/). Previous research has also illustrated that measures used in research assessment have had an effect on scientists' actions (Abbott et al., 2010; Bornmann, 2010). In summary, although the need for multiple indicators was mentioned almost a decade ago (Martin, 1996), with the huge changes in scholarly communication (Veletsianos & Kimmons, 2012), especially in the age of social media, new metrics for capturing the broader impact of research publications beyond bibliometrics are now more widely discussed by information scientists (Cronin & Sugimoto, 2014)

The engagement of researchers with different social web platforms provides a novel opportunity to measure different types of research impact (Cronin, 2013a) and can help to record many kinds of non-scientific research impact (Bornmann, 2014). In particular, social web mentions of scientific publications can be retrieved from various platforms and are often grouped under the umbrella term "altmetrics" (Priem, Piwowar, & Hemminger, 2012). Altmetrics is a new movement that aims to capture new and previously invisible types of impact of scholarly publications on social web platforms such as news sites, Wikipedia, blogs, microblogs, social bookmarking tools and online reference managers (Priem, Taraborelli, Groth, & Neylon, 2011). Altmetrics are not replacement for conventional metrics but complement (Priem, Piwowar, & Hemminger, 2012) current methods such as citation analysis and peer review.

1.2 Research Problem

Although there are many altmetric data sources (Priem, Piwowar, et al., 2012), empirical studies are needed to assess them from different perspectives. Altmetrics is still in early stages and some fundamental issues need to be examined: (a) the suitability and reliability of altmetrics for research evaluation, (b) the identification and reliability of the new metrics, (c) the contexts in which the new data sources can be useful, and (d) differences between different scientific disciplines. This thesis tackles these issues using Mendeley and Faculty of 1000 altmetrics.

Researchers are always looking for ways to understand who is reading their publications (Wouters & Costas, 2012). Amongst Web 2.0 platforms, social bookmarking tools such as CiteULike, Connotea and BibSonomy may help to overcome the lack of global usage data for scientific publications (Haustein & Siebenlist, 2011). A particularly promising example is Mendeley, a free global and collaborative commercial online reference manager tool launched in 2009 that claims to have 2 million users. Online reference management tools are web-based platforms for academics and students to record, manage and share their personal bibliographies. Although there has been much discussion about the value of Mendeley as an altmetrics source (Priem & Hemminger, 2010; Bar-Ilan et al., 2012; Bar-Ilan, 2012), more in-depth research is needed to investigate aspects of using altmetrics in research evaluation.

The combination of peer review and citation analysis has been suggested as a better way to perform research evaluation (Butler, 2007; Moed, 2007). Despite the development of citation databases, there is a lack of a global peer review system for research assessment. Faculty of 1000 (F1000) is a post-publication peer review system for the evaluation of biomedical journal articles that partly exists to fill this gap. Around 10,000 researchers and clinicians across the world evaluate papers on more than 40 subjects within the biomedical domain (F1000, 2012). The recommendations of faculty members in this database may be novel data for research assessment (Waltman & Costas, 2014)In comparison to other altmetrics resources, F1000 provides unique data based on crowd-sourced peer review comments.

1.3 Research Aims and Objectives

The main aim of this thesis is to identify and validate new metrics for research evaluation based on Mendeley and F1000. This aim is guided by the following objectives.

To assess correlations between Mendeley readership counts and citations

This thesis explores whether the relationship between Mendeley readership and citation counts varies across different social sciences and humanities disciplines. Social sciences and humanities are typically not cumulative and topics are not globally agreed on within these disciplines (Becher & Trowler, 2001). Therefore, citation analysis has more limitations for measuring the research performance of these areas than for the hard sciences (Nederhof, 2006). As a result, for last three decades, the development of appropriate indicators for research evaluation of the social sciences and humanities has been important (Moed, Linmans, & Nederhof, 2009). Additionally, “usage metrics” are reasonable measures for fields within the social sciences and humanities because there are many pure readers in these areas (Armbruster, 2008).

To investigate information flows between disciplines based on Mendeley readership

Another task that altmetrics may facilitate is assessing how knowledge travels across disciplinary boundaries, or "information flow". Cross-disciplinary citations are routinely used to measure the information flow from one discipline to another, but this is not ideal (Rinia, Van Leeuwen, Bruins, Van Vuren, & Van Raan, 2002) due to the inherent limitations of citation analysis. For instance, using only citation data to capture information flow across scientific areas will lead to an incomplete analysis as it is based only on publication data, and there are many valid reasons for using scientific papers that will not lead to them being cited. Moreover, citation delays are a common limitation for tracking information flows across different disciplines (Rinia, Van Leeuwen, Bruins, Van Vuren, & Van Raan, 2001). Thus, another objective of this thesis is to examine whether Mendeley can be applied to measure information flows between scientific disciplines.

To identify readers of scholarly publications based on professions

Research articles can be used in different contexts by different types of users both within and outside of academia. To assess the use of research results more effectively, it is important to discover who the readers of scientific publications are and the contexts of their use (Thelwall, 2012). Due to a lack of information about the users of scholarly publications, these issues have not been systematically investigated. Although some case studies using local and small datasets have investigated users of scholarly publications (Niu & Hemminger, 2012; Hemminger, Lu, Vaughan, & Adams, 2007) there are no comprehensive studies at the global level with large-scale datasets. One of the objectives of this thesis is to help to partly fill this gap using the professions of the readers (i.e., professors, researchers, PhD students, undergraduate students, non-academic users) of scientific papers for several disciplines in Mendeley.

To investigate the effect of academic status on the correlations between citations and Mendeley readership counts

Despite correlations between citation counts and Mendeley readers (Li, Thelwall, & Giustini, 2012; Bar-Ilan, 2012) it is not clear to what extent Mendeley readership bookmarking counts actually capture the same or different impacts as citations. Thus, one of the objectives of this thesis is to analyse the effect of academic status on the correlations between citations and Mendeley readership counts across different scientific disciplines.

To establish motivations for bookmarking scholarly publications in Mendeley

Mendeley users can bookmark scholarly records in their personal library for different purposes. Users vary from undergraduate students to professors within academia and other types of users outside of academia. Recently, Mendeley bookmarking counts have been interpreted as indicators of readership but it is unknown to what extent bookmarking counts represent readership. Additionally, the motivations for using Mendeley in general, and the reasons for bookmarking scholarly records in particular, are important issues. In other words, it is important to go beyond numbers and discover what Mendeley readership counts represent. This thesis tackles this problem by investigating motivations for the use of Mendeley in general, and the purposes for bookmarking papers in particular, using qualitative perspectives through a large-scale online survey across different disciplines.

To assess non-standard article impact using F1000 labels

In addition to rating research papers, F1000 members tag articles with predefined labels, including changes to clinical practice, suitability of new drugs, and usability for teaching. They also classify whether new papers confirm, challenge, or reject hypotheses from previous research. F1000 reviewers add value to their evaluations by distinguishing key features of biomedical papers that could represent different dimensions of research. Another objective of this thesis is therefore to assess whether FFa scores[1] can help to identify types of article that are successful but not highly cited.

1.4 Research Contributions

The main novelty of this thesis lies in using new data sources for research evaluation at the article level. This thesis evaluates Mendeley readership counts and F1000 scores across different scientific disciplines with both quantitative and qualitative approaches. This project also explores the potential research impact that can be captured through the proposed metrics. Finally, this thesis provides evidence that the invisible impact of research publications can be revealed through new metrics and can be useful in overcoming some limitations of citations in research evaluation.

1.5 Thesis Structure

This thesis is divided in to six chapters. Chapter 1 describes the background, research problems, aims and objectives. Chapter 2 discusses research related to this thesis, starting with traditional indicators for research evaluation and their limitations. It continues with the metrics based on web resources and discusses the challenges and advantages of altmetrics as well as related case studies. The empirical studies of this thesis are presented in chapters 3, 4, 5 and 6. The research questions, methodologies, findings and discussion for each study are described in each chapter individually. This thesis finishes with conclusions in Chapter 7.

[1] When Faculty Members recommend an article for F1000, they rate the paper as 'Good', 'Very Good' or 'Exceptional' (equivalent to scores of 1, 2 or 3 stars, respectively).

CHAPTER 2: LITERATURE REVIEW

This chapter reviews the literature related to methods and indicators for research evaluation. Section 2.1 introduces the traditional approaches for research evaluation, including citation based-metrics and peer review. The history, development and limitations of these traditional methods are discussed. Finally, literature that compares peer review and citations is reviewed (this is relevant to Chapter 3). Section 2.2 reviews metrics based on web data including webometrics and indicators driven by electronic usage of scientific publications. Section 2.3 summarises the literature related to altmetrics as a new direction for research evaluation, including definitions and concepts, advantages and challenges of altmetrics. Empirical case studies based on altmetric data are also reported.

2.1 Traditional Metrics for Research Evaluation

The impact of research publications can be assessed with both qualitative and quantitative approaches. Citation-based metrics and peer review are quantitative and qualitative methods, respectively, which have been used for research evaluation for many years. In the following sections, these two methods are introduced, and their advantages and challenges are discussed. Related studies that have compared qualitative and quantitative methods are also reviewed.

2.1.1 Peer Review

Peer assessment is a well-established process in academia and plays an important role in modern science (Kostoff, 1997). In this process, researchers invest their time to evaluate the scholarly manuscripts of their peers and to recommend those that meet certain scientific standards. Wouters (1997) used the term “peer review” for different types of peer evaluation in scholarly communication. One of the main applications of peer review in the scholarly ecosystem is reviewing journal articles prior to their publication. This perspective focuses on refereeing activities to identify high quality research for publishing in a journal and mainly supports quality control (Andersen, 2013). Peer judgments are also used for the evaluation of applicants for research fellowships and allocating funding for research proposals (Jayasinghe, Marsh, & Bond, 2003; Bornmann & Daniel, 2005). Some countries also use peer assessment in their national exercises to evaluate research performance and to allocate funding. For instance, the Higher Education Funding Council for England (HEFCE) evaluates the performance of research institutes based on expert opinions (HEFCE, 2011). Peer review is still a valuable way of measuring the quality of research publications but it can be slow and expensive (ABRC, 1990; Smith, 2006).

2.1.2 Citation Analysis

Citation data is a key method used to measure research performance from a quantitative angle. Citation-based metrics have been suggested as a parallel method to peer assessment in research evaluation (Van Raan, 2000). Since the 1960s, citation indexes have been provided by the Institute for Scientific Information (ISI), now Thomson Reuters. These have led to significant changes in the investigation of scholarly communication and research evaluation based on citation analysis. Similarly, Elsevier and Google have established Scopus and Google Scholar with citation indexes.

One of the most well-known indicators based on citation data is the Journal Impact Factor (JIF) of Thomson Reuters. The JIF is published annually. In order to calculate the impact factor for a journal, the total citations that a journal received in the current year to items from the previous two years is divided by the number of citable items (e.g., articles) published in that journal during the last two years. The JIF is popular because of the good reputation of its publisher and its simple definition (Bollen, Sompel, Hagberg, & Chute, 2009). The JIF cannot reflect all of the important dimensions of a journal (Egghe, 1988; Dellavalle, Schilling, Rodriguez, Van de Sompel, & Bollen, 2007) and has been used inappropriately for assessing the research performance of researchers, institutions and research groups (Opthof, 1997; Boell & Wilson, 2010).

For several decades, bibliometrics has used simple analyses based on the number of citations and the number of publications. Gradually, citation-based metrics have been developed for evaluating research output at the level of the journal (e.g. Falagas, Kouranos, Arencibia-Jorge, & Karageorgopoulos, 2008; Leydesdorff, 2009; Moed, 2010). Bollen, Rodriquez, and Sompel (2006), combining the JIF and "weighted PageRank", presented an indicator to reveal the prestige of a journal, as they believed that the JIF shows popularity rather than prestige. At the same time, new metrics, such the h-index (Hirsch, 2005) and the g-index (Egghe, 2006), have been suggested for evaluating research output at the level of the author. Furthermore, due to different citation behaviours in different research fields, new solutions have been proposed to assess performance within academic fields (e.g. Waltman, van Eck, van Leeuwen, Visser, & van Raan, 2011). Multiple indicators based on citation data may provide a wider range of possibilities for scientometricians to evaluate research publications.

There are many publications related to the history, development, opportunities and challenges of citation analysis for research evaluation which are out of the scope of this thesis (see van Raan, 1988; Moed, 2005; Moed, Glänzel, & Schmoch, 2004 for more details).

2.1.2.1 Motivation for Citing

Since researchers cannot cite all previous articles, they need a strategy for selecting their references. Two main theories related to citation behaviour have been developed during the past decades: normative and social constructivist. The former theory assumes that scholars give credit to previous publications via the citations that the new work builds upon (Bornmann & Daniel, 2008) whilst the latter does not. In contrast, social constructivist theory, which is rooted the constructivist sociology of science, criticises the value of citation analysis (Cetina et al., 1991). This theory assumes that citations are not able to measure the academic value of publications because knowledge is developed socially and social contexts affect the citation motivations of scholars (Bornmann & Daniel, 2008). Camacho-Miñano and Núñez-Nickel Manuel (2009) suggested that a multi-layered model is necessary to explain citation practices better. Although Cronin (1984) and MacRoberts and MacRoberts (2010) doubt that precise reasons for citation behaviour can be uncovered, several empirical studies have nonetheless been carried out to establish motivations for citation (see Bornmann & Daniel, 2008). In the medical sciences, an investigation into uncited biomedical articles revealed that the number of authors and references in cited papers was significantly greater than in uncited articles (Stern, 1990). In another study (Hanney et al., 2005), a combination of different variables including locations and reasons for each citation as well as the type of research (basic or clinical) being discussed were used to assess the outcomes of medical publications. Despite the importance of surgery articles, they were cited less in the literature. Features of highly cited and rarely cited articles of The Lancet have been explored based on variables such as the number of authors, references and citations, plus other characteristics related to article type. Articles with more authors and abstract words were cited substantially more (Kostoff, 2007). Recently, Jones et al. (2012) conducted a literature review to extract objective and subjective features of citations in order to develop a new method for citation categorization to assess the quality of individual articles. They claimed that the results of their survey could be useful in measuring the impact of biomedical publications in different generations based on citing publications.

Original research data from an investigation can be reused in new research (Fienberg & Martin, 1985) but some articles with original data are not cited or are cited rarely although their data has been used. This means that some important articles in terms of original data may not be recognized through citation analysis (MacRoberts & MacRoberts, 2010), although the dataset could be considered as a unit for measuring research impact (Sarli & Holmes, 2012). Results of an investigation into a cancer microarray indicated that clinical trials which published their data publicly were cited 70% more than those that did not. This suggests that the availability of research data is an important reason for citation in the medical sciences (Piwowar, Day, & Fridsma, 2007).

2.1.2.2 Challenges of Citation Analysis in Research Evaluation

Since the early days of citation based metrics in research evaluation, there has been a discussion regarding the extent to which these indicators are able to measure research impact. Moreover, using citation analysis in research assessment has some inherent limitations (MacRoberts & MacRoberts, 1989; MacRoberts & MacRoberts, 1996). Citation analysis is able to capture the impact of scholarly publications from the knowledge advancement perspective only (Kostoff, 1998; Merkx, Weijden, Oostveen, Besselaar, & Spaapen, 2007). In other words, citation analysis is limited to the authors' perspectives; however only 15-20% of scholars in the United States have written a journal article (King & Tenopir, 2004). Moreover, a scholarly artefact could be useful in other contexts. For instance, practitioners, undergraduate students (Nicholas et al., 2005), the public (Kurtz & Bollen, 2010) and lecturers use research publications for purposes such as teaching (Kousha & Thelwall, 2008) or professional activities (Schloegl & Stock, 2004), including medical practice (Bennett, Casebeer, Kristofco, & Strasser, 2004; Lewison, 2002). This means that uncited publications may still be useful (Bornmann & Marx, 2014), partly because many non-author professionals also read research articles (Price & Gürsey, 1975; Tenopir & King, 2000). In the same way, science policy makers are looking for different types of impact for research publications within and outside of academia (Drooge, Besselaar, Elsen, & Haas, 2010; De Jong, Van Arensbergen, Daemen, Van der Meulen, & Van den Besselaar, 2011). Additionally, citation metrics are appropriate for the evaluation of theoretical publications but less so for applied research. Thus, there is a worry that a new generation of authors could believe that "citation analysis is a waste of time because authors do not adequately cite those who have influenced their work" (Garfield, 2011, p.2). Another challenge is that citation merely considers formal scholarly communication and new types of communication outside of the conventional form of publishing are ignored (Duin, King, & Van den Besselaar, 2012).

Datasets used in articles are not often cited adequately (MacRoberts & MacRoberts, 2010). This means that the impact of a dataset is overlooked by standard citation indexes. Another challenge is the delay in receiving citations; it varies from 3 months to 2 years or even more for an article to gain a substantial number of citations

(Brody, Harnad, & Carr, 2006). In sum, citation-based indicators are not able to capture the full impact of research, and new ways for discovering more of the full spectrum of research impact are needed.

2.1.3 Peer Review versus Bibliometric Measures

This section compares peer review with bibliometric indicators. This topic is also related to F1000, which is an emerging way of harnessing peer-commentary for metrics. In other words, F1000 is a crowd sourcing way to evaluate the biomedical literature based on expert opinions after publishing, which can be an alternative to conventional peer review in the new era (Huggett, 2012).

Expert judgement is often used as the primary method for evaluating research and its results are sometimes compared to bibliometric measures in order to assess the value of the latter. Maier (2006) used expert opinions and journal impact factors for rating top journals in the field of regional science, finding no significant positive association between the two measures. Peer judgements and the journal impact factors of research articles by Italian researchers in three fields (chemistry, biology and economics) have also been compared, finding a significant but low and medium correlations between the two variables (Reale, Barbara, & Costantini, 2007). Journal impact factors and article citations were also compared to peer judgments in an Italian national research evaluation program, with significant medium correlations being found between both bibliometric indicators and expert ratings (Franceschet & Costantini, 2011).

Several investigations have compared bibliometric indictors and the results of the UK Research Assessment Exercise (RAE), an expert-based method of research evaluation. These studies have tended to find that the two positively correlate.

Oppenheim (1995) compared citation counts and the results of the 1992 UK RAE for Library and Information Science (LIS) departments and found a strong significant correlation between the two measures. Similarly, a strong and significant correlation was observed between UK RAE scores and several bibliometric indicators, including total citations, citations per publication, and citations per academic staff for British LIS departments (Seng & Willett, 1995). Additionally, Smith and Eysenck (2002) investigated the association between citation counts and the UK RAE scores in 1996 and 2001 for British psychology departments and significant and high correlations were reported.

Norris and Oppenheim (2003) compared different citation metrics, such as total and average researchers' citations, with 2001 UK RAE scores for archaeology. The results showed significant and high correlations between citation indicators and UK RAE scores and they suggested that citation measures may be a useful source for the initial rating of research departments. Furthermore, a comparison of the 2001 UK RAE scores and citation measures in the field of music exhibited a strong correlation at the level of departments, while the correlation was weaker for individual citation counts (Oppenheim & Summers, 2008). They suggested that other types of citation data beyond journal articles should be considered for the evaluation of music literature. A large-scale study investigated correlations between citations and UK RAE 2001 scores across a variety of disciplines. In some disciplines of the social sciences, including education, history, sociology, social policy and administration, politics and international studies, there was no correlation between the indicators (Mahdi, D'Este, & Neely, 2008). More recently, Kousha, Thelwall, and Rezaie, (2011) compared citations to books submitted to the 2008 UK RAE and expert judgements for seven social sciences and humanities disciplines and found weak correlations between these measures. In summary, the relationship between UK RAE scores and citations differed across disciplines, which is presumably due to differing citation behaviours between academic fields.

In an earlier study, the performance of six research groups was measured using both peer judgements and citation analysis, with the results also being compared. Although each method had its own unique features to assess research performance, correlations were reported between the results of expert assessment and citations (Nederhof & van Raan, 1993). In another investigation, the relationships between various citation-based indicators and expert ratings for different research groups for a university in Norway were measured. There was a positive but not strong correlation between expert ratings and all the bibliometric indicators. The authors argued that citation metrics could not be an absolute alternative for peer evaluation but only a complementary indicator (Aksnes & Taxt, 2004). Van Raan (2006) compared the h-index and the crown indicator (CPP/FCSm[2]), both citation-based, with expert evaluations of chemistry research groups in the Netherlands. The quantitative and qualitative indicators correlated very well; however, the new crown indicator was more suitable for smaller research groups with "less heavy citation traffic". Opthof and Leydesdorff (2011) challenged the new crown indicator. They used previously-released data (van Raan, 2006) to re-examine correlations between citation parameters and peer review results for chemistry research groups in the Netherlands. The qualitative indicator and the two citation metrics by the CWTS (CPP/FCSm) had no significant correlation. Waltman et al. (2011) believed that Opthof and Leydesdorff (2011) used a statistical method that only shows "presence and the

[2] The CPP/FCSm is an indicator developed by The Centre for Science and Technology Studies (CWTS) with the aim of normalization of citation in different fields. It has been renamed the Crown indicator (see http://arxiv.org/pdf/1003.2167.pdf).

absence of a relation” rather than determining the strength of the relationship between the two variables. Thus, Waltman et al. (2011) investigated a larger amount of data with a more suitable statistical technique to measure the degree of relation between the two indicators and revealed that the CPP/ FCSm parameters correlated significantly with expert opinions. The CPP/FCSm indicator has been renamed as the MNCS (Mean Normalized Citation Score) which is part of The CWTS Leiden Ranking system (see http://www.leidenranking.com/methodology/indicators).

2.2. Online Metrics for Research Evaluation

The advent of the Web provided opportunities for scholars and publishers to put academic publications online in electronic formats. As a result, the publication and use of electronic information resources by scholars has offered new types of data for research evaluation. The following sections review literature related to webometrics and electronic usage metrics.

2.2.1 Webometrics

Informetrics is the study of any form of information using quantitative approaches (Tague-Sutcliffe, 1992), including bibliometrics, scientometrics and citation analysis (Egghe & Rousseau, 1990). Cybermetrics, on the other hand, is the study of electronic information on the Internet using bibliometric methods (Björneborn, 2004). In the late 1990s, the growth of scholarly information on the web provided an opportunity to measure new types of scholarly communication. Thus, the web is a medium through which the diverse impacts of scholars could be identified (Cronin, Snyder, Rosenbaum, Martinson, & Callahan, 1998). Because of this and other observations about the potential of the web for bibliometrics (e.g., Almind & Ingwersen, 1997), the new area of webometrics emerged a sub-field of cybermetrics and informetrics for the study of web-based phenomena drawing on bibliometric methods (Björneborn & Ingwersen, 2004). The web impact factor using hyperlink data was one of the early webometrics indicators and was based on the traditional JIF (Ingwersen, 1998). Several case studies found correlations between the numbers of links to their websites and the research performance of universities (Smith, 1998; Vaughan & Thelwall, 2003; Qiu, Chen, & Wang, 2004), departments (Li, Thelwall, Wilkinson, & Musgrove, 2005) and research groups (Barjak & Thelwall, 2008). At the level of individual journals, significant correlations between the JIF and the web impact factor were found for journals in specific disciplines (Vaughan & Hysen, 2002). Disciplinary differences and the age and content of a journal website are important factors for correlations between the two measures (Vaughan & Thelwall, 2003).

In order to measure the impact of research articles in different online resources, "web citation" counting has been suggested (Vaughan & Shaw, 2003). With this method, the appearance of a specific article can be tracked using search engines. Several empirical studies found significant correlations between web citations and traditional citations (Vaughan & Shaw, 2003; Vaughan & Shaw, 2005; Kousha & Thelwall, 2007a). However, a qualitative analysis revealed that the majority of Google web citations appeared in non-scholarly contexts rather than in the references of documents (Kousha & Thelwall, 2007b).

The stability of webometric methods is mainly dependent on the policies of commercial search engines. For instance, the stopping of hyperlink searches by commercial search engines was a serious barrier in this area, although URL citation searches can be a partial replacement (Thelwall, 2011).

Webometric studies have developed new techniques to measure research impact beyond conventional citation data and suggested new types of research impact based on web data, including teaching (Kousha & Thelwall, 2008) and online presentation impact (Thelwall & Kousha, 2008). Additionally, Kousha and Thelwall (2009) have suggested Google Books as a source to track the citation impact of books, which was previously a problem for measuring research impact in book-based disciplines in the social sciences and humanities (Cronin, Snyder, & Atkins, 1997; Cronin et al., 1997; Nederhof, 2006). New forms of scientific documents have also been examined using webometric techniques. For instance, Kousha, Thelwall and Rezaie (2010) provided evidence that the educational impact of, or public interest in, online scholarly images can be identified through the Web. Several publications have provided comprehensive reviews of webometrics (e.g. Thelwall, 2012; Kousha & Thelwall, 2014).

2.2.2 **Electronic Usage Statistics**

The publishing of scholarly documents in electronic format and advances in collecting large-scale usage data during the last decade have provided new opportunities for in-depth analyses of the information seeking behaviour of scholars. As a result, usage indicators have become important (Harnad, 2008).

From the research evaluation point of view, usage data may help to measure scientific impact and to supplement citation analysis (Bollen, Van de Sompel, Smith, & Luce, 2005; Ian Rowlands & Nicholas, 2005; Schloegl & Gorraiz, 2011). Concepts related to journal usage metrics have appeared, with different terms such as “readership”, “usage,” and “downloads” in the literature (Kurtz & Bollen, 2010). Journal usage metrics refer to indicators based on the usage data of electronic journals (Rowlands & Nicholas, 2007) that provide a reasonable evaluation of the journals (Hahn & Faulkner, 2002), such as downloads or online views. Similarly, readership has been defined as “full-text downloads” (Haque & Ginsparg, 2009) or online views of a particular paper (Kurtz et al., 2005).

Researchers in information science have used traditional citation methods to develop new indicators based on download data. Several investigations have compared usage statistics for journal articles with citation counts. Kurtz et al. (2000) found moderate correlations between readership and citation for astrophysics publications. Brody, Harnad, and Carr (2006) found correlations between citations and download counts for different disciplines in arXiv. Usage counts of the top 200 downloaded papers from RePEc, a repository in economics, have been compared with their citation measures from WoS and Google and a moderate association between the number of citations and downloads was reported (Chu, Krichel, & Blvd, 2007). Moed (2005, p.1906) explored usage data and citations during the early months after publication. He came to the conclusion that “initial downloads” and citations relate to distinct phases in the process of collecting and processing relevant scientific information that eventually leads to the publication of a journal article”. Thelwall (2012) also argued that there is enough evidence to rely on usage data as a scientometric factor. Nevertheless, it is not easy to make a direct and general conclusion about the relationship between citation and usage data due to a variety of factors interfering in their association. For instance, citation behaviours vary across disciplines (Moed, Leeuwen, & Reedijk, 1998) as do usage practices in different fields (Rousseau, 2000;

Tenopir, 2003). The role of academics and types of publication can also affect the relationship between usage and citation practices (Kurtz & Bollen, 2010).

In addition to measuring readership counts through usage, information corresponding to the geography and affiliations of readers may be accessible (Kurtz, Eichhorn, Accomazzi, Grant, Demleitner, Murray, et al., 2005) for bibliometric research. Although much research has anaysed usage data from different points of view, only the basic concepts of readership have been investigated to date (Nicholas et al., 2005; Jansen, 2006; Mayr, 2006; Bollen & Sompel, 2008).

Using local and global citation and usage data is an important issue (Bollen & Sompel, 2008) that can be applied to reveal different patterns. Several studies have analysed local usage data for different academic institutions (Duy & Vaughan, 2006; McDonald, 2007; Bollen & Sompel, 2008). Some investigations used both local citations and local usage simultaneously (Blecic, 1999; Sridhar, 1990; McDonald, 2007). In contrast, global usage data is not limited to users' behaviour within an academic organization or university (e.g., an academic field of science) (Schloegl & Gorraiz, 2011). Due to confidentiality and marketing aspects of user logs, global usage data from commercial publishers is rarely available and researchers have used scholarly repositories (e.g., arXiv.org) and open access journals because of the availability of their data (Schloegl & Gorraiz, 2010).

At the level of journals, Bollen and Sompel (2008) defined the "Usage Impact Factor" (UIF) for journal evaluation. The Download immediacy index (Wan, Hua, Rousseau, & Sun, 2010,) and usage half-life (Rowlands & Nicholas, 2007) have also been suggested as new indictors. The download immediacy index is "the number of downloads of a journal's articles within one publication year, divided by the number of published articles by that journal in that same year" (Wan et al. 2007). Rowlands and Nicholas (2007) operationalized usage half-life as the median age of all publications of particular journal that were downloaded in a specific year without considering their age. Several empirical studies have compared usage data with the JIF. For instance, a comparison between usage data for Rouen University Hospital digital library and JIFs showed a significant but low correlation (Darmoni & Roussel, 2002). Wulff and Nixon (2004) found that usage of both print and electronic articles had low correlations with the JIF. Duy and Vaughan (2006) reported no correlation between downloads and the JIF for chemistry and biochemistry journals at Concordia University. Additionally, Bollen and Sompel (2008) found negative correlations between the UIF and the JIF based on the use of California State University library for all disciplines except education. Low and insignificant correlations between usage metrics and JIFs suggest that these indicators reflect different aspects of research impact.

In sum, usage statistics are able to capture broader research activities than are citations (Kurtz & Bollen, 2010) and are obtainable earlier than citation indicators (Brody, Harnad, & Carr, 2006). However, usage-based metrics have used local usage data since global usage statistics are hidden by commercial publishers (Schloegl & Gorraiz, 2010) because of privacy and marketing concerns. The value of a download also depends on who accessed an article and how it was used (Thelwall, 2012). Moreover, the availability of an article through multiple platforms (Rowlands & Nicholas, 2007) and data aggregation are other challenges for obtaining accurate usage data (Haustein & Siebenlist, 2011).

2.3 Altmetrics: Definition and Concepts

Scholars can now communicate in the social web, including social bookmarking sites, Twitter, blogs, and wikis. The engagement of researchers with different social web platforms provides a novel opportunity to measure different types of research impact (Cronin, 2013a) and can help to capture many kinds of non-scientific research impact (Bornmann, 2014). In particular, social web mentions of scientific publications can be retrieved from various online platforms and are often grouped under the umbrella term *altmetrics* (Priem, Piwowar, & Hemminger, 2012). The altmetrics movement was formed with the aim of capturing new and previously invisible impacts of scholarly publications based on crowdsourced data in social web platforms such as blogs, microblogs, social bookmarking tools and online reference managers (Priem, Taraborelli, Groth, & Neylon, 2011). Altmetrics invokes non-conventional metrics for discovering new types of research impact (Liu & Adie, 2013). The Public Library of Science (PLoS) has also suggested a series of metrics including citations, usage data, and altmetrics in one "article-level metrics"[3] package that aims to provide transparent data and metrics for individual articles (Fenner, 2013). Altmetrics and article-level metrics may have common borders but they differ in several ways. Article-level metrics cover citations, usage statistics and new measures, while Altmetrics deal with new metrics based social web data (Fenner, 2014). Altmetrics is as a sub-area of both scientometrics and webometrics and is close to webometrics but it uses social web platforms as data sources instead of the general web (Priem, Groth, & Taraborelli, 2012). Rousseau and Ye (2013) argued that although the idea behind altmetrics is valuable, the term is not appropriate. They

[3] http://article-level-metrics.plos.org/alm-info

suggested “influmetrics” instead, as a sub-division of webometrics. Similarly, Cronin (2013b) argued that “complementary” is a better term than “alternative” in this context since the latter term has the commutation that altmetrics might replace citation metrics.

2.3.1 Advantages of Altmetrics

Altmetrics can have several advantages. Due to the time-delay in citations, real-time data is only possible with altmetrics or usage data. Data collection for altmetrics is based on open Applications Programming Interfaces (APIs) (Priem, Piwowar, & Hemminger, 2012) which are more accessible than usage data (Priem et al., 2011). The accessibility of data through open APIs has two advantages, openness and transparency, that can help to overcome dependency on commercial publishers (Wouters & Costas, 2012). Unlike usage data, the integration of different data sources from social platforms is easy with open APIs (Priem, 2014). Reproducibility, the ability to replicate and validate research, is a principle in scholarly methods that can be a new aspect in the evaluating of research publications (Iorns, 2012). The easy integration of data into altmetrics resources is a good opportunity to provide an infrastructure for validating reproduced findings. Hence, Mendeley, Figshare, Science Exchange and PLOS have started a joint initiative for reproducibility[4] in research. Altmetrics data are not restricted to the author community, and it is possible to track the use of scholarly publication in different contexts within and outside academia. Additionally, due to the openness of altmetrics data, detailed information

[4] https://www.scienceexchange.com/reproducibility

about users of scholarly publications can be available. For example, Mendeley provides the academic statuses, countries and disciplines of some readers. Thus, it is possible to study the use of science publications beyond citations.

2.3.2 Implementing Altmetrics Data

As mentioned before, a single indicator or even several metrics obtained from similar data types are not able to capture the all aspects of research impact. In theory, altmetrics have the potential to identify research impacts at different levels and different contexts but good, clean, systematic data and metrics are needed to prove this claim. As a result, several initiatives have focused on defining methods for capturing and providing altmetrics. The main tools that aggregate altmetrics data and their multi-dimensional perspectives in research evaluation are discussed briefly in the following paragraphs. For an extensive review of altmetrics tools and their features, see Wouters and Costas (2012).

PLOS is one of the pioneers in providing new data sources and citations for research evaluation at the article level. Based on several types of aggregated data, PLOS organized research impact into several themes: “viewed”, “saved”, “discussed”, “recommended” and “cited” (Fenner, 2013). Figure 2.1 illustrates the themes and data sources together in a framework.

Figure 2.1: PLOS article level classifications. (Lin & Fenner, 2013)

ImpactStory (formerly Total Impact) is an open-source and online platform which was established in 2011. The aim of this tool is to reveal diverse impacts of research articles' impacts via different types of data sources[5]. Figure 2.2 shows ImpactStory's perspectives about data sources and research impacts for scholarly publications.

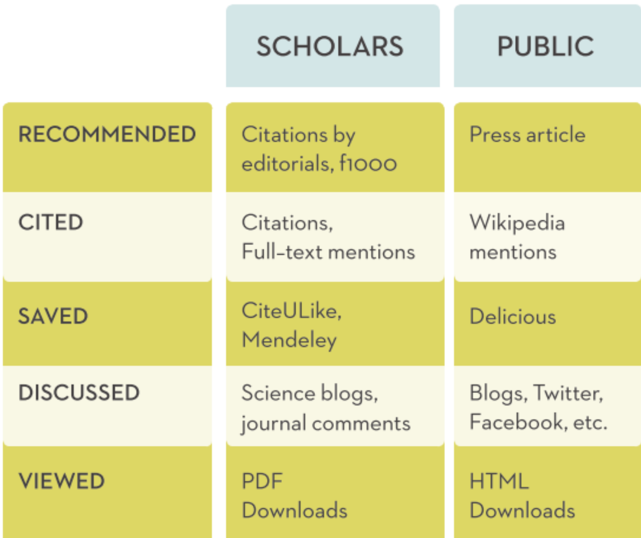

Figure 2.2: The thematic structure of altmetrics from ImpactStory. (Priem, 2013)

[5] http://impactstory.org/about

In late 2011, Plum™ Analytics[6] (recently acquired by EBSCO[7]) published several metrics for researchers and funders. Plum™ Analytics classifies research impacts into four main categories: usage, captures, mentions, social media and citations (Figure 2.3).

Category	Metrics	Example Sources*
Usage	Downloads Views Book Holdings ILL Document Delivery	PLOS WorldCat ePrints Vimeo dSpace
Captures	Favorites Bookmarks Saves Readers Groups Watchers	CiteULike Slideshare Github Mendeley YouTube
Mentions	Blog Posts News Stories Wikipedia Articles Comments Reviews	Wikipedia Facebook SourceForge Reddit
Social Media	Tweets +1s Likes Shares Ratings	Facebook Twitter Google Plus
Citations	Citation Count	Pubmed Scopus

Figure 2.3: Plum™ Analytics classifications. (Buschman & Michalek, 2013).

Altmetric.com (http://www.altmetric.com) was established as a start-up in London in 2011. Altmetric.com aims to capture and analyse activities related to scientific

[6] http://www.plumanalytics.com
[7] http://www.ebscohost.com/newsroom/stories/plum-analytics-becomes-part-of-ebsco-information-services

publications in online settings, such as Twitter, blogs, Mendeley, CiteUlike, Facebook, Google Plus and news platforms. These activities are summarised in a colourful donut. As shown in Figure 4.4, statistical data about each online activity is displayed using a different colour and the overall Altmetric score (for details of the calculation, see http://goo.gl/Ukql1x) is in the middle of the donut. However, Altmetric.com focuses on providing clean and accurate data rather than on the evaluation of research publications (Liu & Adie, 2013). Several publishers and journals have integrated their system with Altmetric.com to monitor the performance of their publications.

Figure 2.4: An example of online activity surrounding an article, provided by Altmetric.com.

In summary, altmetrics tools and related infrastructures have been developed successfully; however, altmetrics are still at an early stage of development. All of the aforementioned tools offer altmetrics but each tool provides different data. For example, Plum™ Analytics collects data from interlibrary loans and holdings of books in libraries, which are not provided by others. The services that the altmetrics tools offer are also different. For instance, the APIs of Altmetric.com and PLOS have fewer limitations in comparison to the ImpactStory and Plum™ Analytics APIs.

2.3.3 Challenges of Altmetrics

Altmetrics as a research area is still in its early stages, and motivations behind the numbers are unknown. In other words, the numbers obtained from different sources need to be contextualized (Wouters & Costas, 2012; Konkiel, 2013) so that it is clear whether they reflect academic interest, public interest, spam or something else. Although one of the purposes of altmetrics is measuring research impact beyond academia, it is not easy to determine scholarly and non-scholarly audiences in different platforms (Haustein, 2013). Unlike the traditional indicators, which use the scholarly literature, altmetrics rely on new media that have a more dynamic nature; thus, inconsistency of data is another limitation (Fenner, 2014). Additionally, the durability of data and platforms is another challenge (Liu & Adie, 2013). The potential for manipulating and gaming altmetrics data is also a serious limitation (Priem, Parra, Piwowar, Groth, & Waagmeester, 2012) which is rooted in the lack of quality control on the social web. The majority of new metrics are more appropriate for recent publications and less suitable for old papers. Although the National

Information Standards Organization has started to develop standards for altmetrics[8], they are not yet available (Konkiel, 2013). Additionally, altmetrics are prone to biases towards scholars with more Web visibility, who are mainly younger (Priem, 2014). Finally, the behaviours of scholars in social media are not similar across disciplines, countries and languages, and therefore the normalization of altmetrics for different contexts needs to be considered (Wouters & Costas, 2012).

2.3.4 The Availability of Data for Altmetrics

Around ten years ago, Cronin (2005, p.196) predicted that "there will soon be a critical mass of web-based digital objects and usage statistics on which to model scholars' communication behaviors—publishing, posting, blogging, scanning, reading, downloading, glossing, linking, citing, recommending, acknowledging—and with which to track their scholarly influence and impact, broadly conceived and broadly felt". In practice, several new online platforms have been launched on the Web for facilitating scholarly communication in new forms with different purposes. These include including social bookmarking sites, post-publishing recommendation platforms, scientific blogs, and Twitter. Social websites and their facilities for scholars have been discussed before (see Wouters & Costas, 2012; Priem & Hemminger, 2010). These tools are potential sources of data for measuring the impact of scholarly publications at different levels, including articles, organizations and researchers. Online reference managers and social bookmarking sites are personal information management tools that researchers can use in their daily activities. User interactions in social bookmarking sites can provide valuable data that could be useful for research evaluation (Neylon & Wu, 2009). For example, these social platforms provide opportunities to trace the global usage of scientific

[8] http://www.niso.org/news/pr/view?item_key=72efc1097d4caf7b7b5bdf9c54a165818399ec86

publications (Haustein et al., 2010). There seems to be plentiful data about biomedicine articles in social bookmarking platforms (Priem et al., 2012). A case study of BibSonomy revealed that the most bookmarked publication types were journal articles (Borrego & Fry, 2012). Priem, Piwowar and Hemminger (2012) explored a large sample of papers published by the Public Library of Science (PLoS). Around 80% of the PLoS articles were covered by Mendeley while 31% and 10% of these papers were bookmarked on CiteULike and Delicious, respectively, although it is not completely fair to compare statistics between the sites because they are used and record information in different ways. Around 10% to 12% of the sample were tweeted or mentioned on Facebook and less than 10% of the papers were cited in blogs or reviewed by Faculty of 1000 (F1000, now F1000Prime). Similarly, previous studies have reported that the coverage of Mendeley is more extensive than that of CiteULike for a sample of articles published in Science and Nature (Li, Thelwall, & Giustini, 2012), with similar results being found for publications in the field of bibliometrics (Haustein, Peters, Bar-Ilan, et al., 2013). It has also been reported that Mendeley had the highest coverage among altmetrics resources for 20,000 random publications indexed in WoS (Zahedi, Costas, & Wouters, 2013).

In contrast, analysing the entire F1000 database, Waltman and Costas (2014) discovered that as few as 2% of biomedicine articles had been reviewed by F1000 experts. Again, the figures are not directly comparable because F1000 articles are reviewed whereas Mendeley articles are only recorded on the site. A large-scale study of PubMed articles in 11 social media resources (excluding Mendeley) reported that less than 20% of the papers were covered by most of the resources (Thelwall, Haustein, Larivière, & Sugimoto, 2013), with Twitter having the most extensive coverage at less than 10% for 2010 to 2012 PubMed articles and reviews

(Haustein, Peters, Sugimoto, Thelwall, & Larivière, 2013a). In another large-scale and multidisciplinary study, Costas, Zahedi, and Wouters (2014) discovered that research papers had higher coverage (13.3%) in Twitter compared to other social websites, including Facebook walls, blogs, Google+ and news outlets. A later Mendeley analysis of the same set of 1.4 million PubMed papers reports that 66% had at least one Mendeley reader (Haustein et al., submitted). Finally, a survey of bibliometricians reported that most had LinkedIn profiles (68%) and around half had Twitter accounts while 20% were users of Academia.edu, Mendeley, and ResearchGate (Haustein, Peters, Bar-Ilan, et al., 2013).

2.3.5 Correlations between Altmetrics and Citations

Several studies have assessed the value of altmetric indicators by comparing them with traditional metrics. For example, several bookmarking-based metrics and some traditional indicators have been compared for physics journals (Haustein & Siebenlist, 2011). Significant positive correlations between bookmarking data from CiteUlike and Mendeley and citations for a sample of articles published in Science and Nature provide evidence for the value of bookmarking data for research evaluation (Li et al., 2012). Nevertheless, based upon moderate correlations between Mendeley readership and citation counts for articles, Bar-Ilan (2012), Haustein et al. (2013), Haustein et al. (submitted) and Zahedi et al., (2013) all concluded that reading and citing are not similar scholarly activities. Li and Thelwall (2012) found positive correlations between Mendeley readership counts and bibliometric indicators for a sample of papers in the field of genomics and genetics. In a large-scale study, moderate positive correlations between Mendeley readership counts and citations were found for PubMed papers and it was concluded that readership and citations reflect different types of research impact (Haustein et al., submitted). Using Mendeley data, Thelwall & Maflahi (2014) discovered that readers of scientific papers tend to have the same country affiliation as the authors.

In terms of predicating future citation counts, Twitter mentions of articles in a single open access online medical informatics journal could predict future citations (Eysenbach, 2011). Similarly, Shema, Bar-Ilan, and Thelwall (in press) reported that papers mentioned in science blogs received more citations later. A large-scale analysis of biomedicine documents (Haustein et al. 2013) found weak correlations between citation counts and Twitter mentions, and concluded that these two metrics reflect different types of impact. Similarly, low correlations between citations and several social media metrics including Twitter, Facebook, blogs and news mentioned have been reported (Thelwall et al., 2013; Costas et al., 2014). Based upon insignificant correlations between indicators derived from the academic social website Academia.edu and bibliometric indicators, Thelwall & Kousha (in press) concluded that the informal scholarly communications in Academia.edu probably do not reflect traditional academic impact or prestige, however.

Another investigation found a positive correlation between Wellcome Trust reviewers' assessments and F1000 ratings (Allen, Jones, Dolby, Lynn, & Walport, 2009). Wardle (2010) compared F1000 rates and citations for articles in seven important ecological journals. The results showed that 46% of all publications that were highlighted by F1000 reviewers had not been highly cited articles. Comparing F1000 article factors (FFa) and several bibliometric measures provided by InCites Thomson Reuter, Bornmann and Leydesdorff (2012) found the highest correlation between the FFa scores and “Percentile in Subject Area”. Li and Thelwall (2012) compared the FFa scores of 1,397 reviewed articles in the fields of Genomics &

Genetics with citations and the journal impact factor, and reported significant correlations between the FFa scores and both citation metrics. Recently, Waltman and Costas (2014) compared all F1000 recommendations with citations and found significant but weak correlations between citation counts and F1000 recommendations. The empirical studies that compared altmetrics with citation-based metrics are summarized in Table 2.1.

Table 2.1. Empirical studies comparing altmetrics and citation-based metrics.

Metrics	**Sample**	**Correlation**	**Paper**
Mendeley readership/citations	Nature and Science articles 2007	0.559** Spearman	Li, Thelwall, and Giustini (2012)
	PLoS journals articles 2003-2010	0.300/ 0.400, 0.500 Spearman	Priem, Piwowar and Hemminger (2012)
	JASIST articles 2001-2010	0.500	Bar-Ilan (2012)
	A random sample of 20,000 publications with DOIs (published from 2005 to 2011)	0.307 Spearman	Zahedi et al. (2013)
	Genomics and Genetics articles published in 2008 selected by F1000 Faculty	0.686**/WoS 0.682**/Scopus 0.694**/Google Scholar Spearman	Li and Thelwall (2012)
Delicious bookmarks/citations	a random sample of 20,000 publications with DOIs (published from 2005 to 2011)	0.002 Spearman	Zahedi et al. (2013)
CiteULike bookmarks/ citations	Genomics and Genetics articles published in 2008, selected by F1000 Faculty	0.213/WoS 0.231/ Scopus 0.242/ Google Scholar Spearman	Li and Thelwall (2012)
	A random sample of 20,000 publications with DOIs (published from 2005 to 2011)	0.119 Spearman	Zahedi et al. (2013)
Facebook mentions/ citations	PubMed articles with at least one altmetric mention	0.050** Spearman	Thelwall et al. (2013)

Metrics	Sample	Correlation	Paper
	A random sample of 20,000 publications with DOIs (published from 2005 to 2011)	-0.005 Spearman	Zahedi et al. (2013)
	All articles covered by Altmetric.com	0.099 Spearman	Costas et al., (2014)
PLoS-hosted comments/ citations	PubMed articles with at least one altmetric mention	0.034** Spearman	Thelwall et al. (2013)
Google+/citations	All articles covered by Altmetric.com	0.060 Spearman	Costas et al., (2014)
	PubMed articles with at least one altmetric mention	0.201** Spearman	Thelwall et al. (2013)
Blogs/citations	Altmetric.com	0.126 Spearman	Costas et al. (2014)
Page views and shares/citations	Journal of Medical Internet Research Articles 2009/2010	0.510* Spearman	Eysenbach (2011)
Tweets/citations	PubMed articles with at least one altmetric mention	-0.190 Spearman	Thelwall et al. (2013)
	All articles covered by Altmetric.com	0.167 Spearman	Costas et al. (2014)
	biomedical literature (from WoS and PubMed) for the 2010 to 2012	0.114** Spearman	Haustein, Peters, Sugimoto, et al. (2013b)
	philosophers in Academia.edu	0.116 Spearman	Thelwall and Kousha (2014)
Profile views in adcademia.edu/ citations	philosophers in Academia.edu	0.122 Spearman	Thelwall and Kousha (2014)
Profile views in adcademia.edu/ citations	philosophers in Academia.edu	0.058 Spearman	Thelwall and Kousha (2014)
Document views in adcademia.edu/ citations	PubMed articles with at least one altmetric mention	0.373** Spearman	Thelwall et al. (2013)
RH†/citations	PubMed articles with at least one altmetric mention	0.088** Spearman	Thelwall et al. (2013)
MSM††/citations	All articles covered by Altmetric.com	0.076 Spearman	Costas et al. (2014)
News/citations	PubMed articles with at least one altmetric mention	0.062** Spearman	Thelwall et al. (2013)
Reddits/citations	PubMed articles with at least one altmetric mention	0.033** Spearman	Thelwall et al. (2013)
Forums/citations	PubMed articles with at least one altmetric mention	0.048** Spearman	Thelwall et al. (2013)
Q&A†††/citations	PubMed articles with at least one altmetric mention	0.005** Spearman	Thelwall et al. (2013)
Pinners††††/citations	PubMed articles with at least one altmetric	0.009** Spearman	Thelwall et al. (2013)

Metrics	Sample	Correlation	Paper
	mention		
F1000 recommendations /citations	Genomics and Genetics articles published in 2008 selected by F1000 Faculty	0.303** Spearman	Li and Thelwall (2012)
F1000 recommendations /citations	a random sample of 20,000 publications with DOIs (published from 2005 to 2011)	0.097 Spearman	Zahedi et al. (2013)

† Research highlights are from Nature Publishing Group journals
††The Mainstream Media citation
††† Online questions and answers
†††† Pinterest.com
(* = statistically significant at the 5% level, ** = statistically significant at the 1% level)

To summarize, although many metrics based on social web platforms correlate with bibliometric indicators for specific sets of articles, the correlations are weak or moderate. As shown in Table 2.1, Mendeley readership counts had the highest correlations among all measures. Evaluating correlations has been suggested as the first step in the evaluation of altmetrics (Sud & Thelwall, 2014). Hence, further studies are needed to discover different aspects of these new metrics.

2.4 Theoretical Background

2.4.1 Interdisciplinary Knowledge Transfer Using Citation Data

Science policymakers and funders sometimes promote interdisciplinary research to overcome sophisticated research problems (Levitt & Thelwall, 2011) and cross-fertilization seems also to be a vital element in modern science (Morillo, Bordons, & Gómez, 2003). Thus, researchers may use more publications from outside their disciplines (Bordons, Morillo, & Gómez, 2005). Interdisciplinarity can be conceptualized in two different ways: big and small (Rinia, 2007). Small interdisciplinarity deals with interactions between sub-disciplines while big interdisciplinarity refers to relations between different disciplines. It seems that some disciplines are mainly “donors” while others are “receptors” (Pair, 1980). In other words, some disciplines provide ideas, methods or findings for other disciplines while several disciplines are mainly importers of ideas from others. It is therefore increasingly important to study the information flow between disciplines to uncover the contributions of scientific disciplines from different points of view.

A number of previous studies have attempted to measure interdisciplinarity in the social sciences and humanities mainly based on citation analysis. Urata (1990) used citation flows and expert migration to identify relationships between social sciences and humanities disciplines in Japan. The results revealed that sociology and education imported many ideas from other disciplines while psychology, linguistics, philosophy and history exported though to other areas. For the social sciences, Larivière and Gingras (2010) found that interdisciplinary decreased from 1965 to 1992, but rose sharply after 1994. Levitt and Thelwall (2011) investigated interdisciplinarity changes in social sciences disciplines in 1990 and 2000 with similar results: interdisciplinarity diminished between 1980 and 1990 but increased strongly from 1990 to 2000. Stevens (1990) examined the relationship between planning and other social sciences disciplines. He found that half of the planning information came from economics while geography, environmental studies and economics were the main users of planning publications. An investigation into articles from four major sociology and political science journals indicated that the boundaries of these disciplines were not limited (Pierce, 1999). Goldstone and Leydesdorff (2006) claimed that cognitive science, as an interdisciplinary subject, is a hub for knowledge exchange between computer science, neuroscience, psychology and education. Cognitive science articles were often used by computer scientists while cognitive science researchers cited psychology publications more.

Neeley (1981) applied citation analysis to measure the relationship of management to other social science fields, finding that management scholars often cited other disciplines but not vice versa. Another study of management journals revealed that this field was a significant donor for psychology while a large amount of information was imported from economics, psychology, and sociology (Lockett & McWilliams, 2005).

Cronin and Pearson (1990) analysed citations to the scholarly artefacts of some senior information scientists and found that few of these publications were used by scholars from outside of the field. Conversely, an empirical study in 2005 showed that the pattern of LIS (Library and Information Science) research had changed in terms of interdisciplinarity and LIS articles had been cited by several other disciplines (Tang, 2005). Cronin and Meho (2007) used large-scale data to re-examine the conclusions of Cronin and Pearson (1990), finding that information science had transferred ideas to other disciplines more and used publications from computer science, engineering, and business and management more in the last decade. More recently, LIS has had the highest increase in interdisciplinarity among the other social sciences disciplines (Levitt & Thelwall, 2011).

2.4.2 Knowledge Flows in Bookmarking Data

Social bookmarking and reference data have also been used for discovering relations between disciplines. Jiang, He, and Ni, (2011) clustered journals and authors from occurrences and co-occurrences in CiteULike and compared the results with clusters generated from cross-citations and co-citations from WoS. The results indicated that although CiteULike data was able to show relationships between publications, it is not as good as citation-based data for showing connections in the literature because of data sparseness. Kraker, Körner, Jack, and Granitzer (2012) used “co-readership” of saved articles in Mendeley users’ profiles to discover emerging areas in the technology-enhanced learning field. An analysis of 1,025 personal libraries in Mendeley revealed that it was possible to uncover emerging topics based on usage data. Although the sample investigated was limited to a small subject, they concluded that Mendeley could be useful for real time and accurate knowledge domain visualization.

In summary, measuring information flows based on citation analysis is an established method for discovering disciplines’ influences on each other. Nevertheless, it is restricted to citation data and hence the behaviour of authors so it may not reveal all knowledge transfers between scientific disciplines. Although some studies have shown that social bookmarking data could be valuable for measuring connections between disciplines, they were restricted to CiteUlike with small database in comparison with Mendeley (Jiang et al., 2011) or were limited to a particular research area (Kraker, Körner, Jack, & Granitzer, 2012). Additionally, characteristics of scientific disciplines in terms on information flow based on social bookmarking data have not been addressed clearly in previous research. Thus, a systematic study for discovering information flows based on social bookmarking data particularly for social sciences and humanities with many readers is needed.

2.4.3 Professions and Science

The responsibilities of different professions and the status of academics can both affect the roles and contexts in which individuals use scholarly publications. For example, younger researchers read more papers (Tenopir, King, Spencer, & Wu, 2009) and cite more resources in their publications (Pancheshnikov, 2007; Barnett & Fink, 2008; Larivière, Sugimoto, & Bergeron, 2013) in comparison to senior researchers. Niu and Hemminger (2012) found academic status to be an important issue in information seeking behaviour for faculty members, students and staff at five US universities. Interviewing scholars in the field of humanities, Ge (2010) revealed that PhD students and assistant professors use electronic resources more than do associate professors and professors. Jamali and Nicholas (2006) found that PhD students browse electronic journals more than do senior scholars in physics and astronomy. Catalano (2013) concluded that although Masters and PhD students both use the web for information searching, the latter believe that references provided by faculty members are more reliable. Whitmire (2002) argued that the information seeking behaviour of undergraduate students could be different from that of graduate students and faculty members but they can have similar information seeking behaviour because some students used resources suggested by faculty members more than other references (Korobili, Malliari, & Zapounidou, 2011).

Outside academia, practitioners and developers use research publications in their daily activities (Bollen & Van De Sompel, 2008) but the roles, tasks and the contexts in which they use information can affect their information seeking behaviours (Leckie, Pettigrew, & Sylvain, 1996). For example, a survey of non-author physicians in Canada discovered that 73% of the participants read journal articles (McAlister, Graham, Karr, & Laupacis, 1999). Another study reported that journal articles were the most useful publication type to fulfil the information needs of residents in a hospital (Schilling, Steiner, Lundahl, & Anderson, 2005). The information seeking habits of engineers in different fields (Ellis & Haugan., 1997; Kwasitsu, 2003; Freund, Toms, & Waterhouse, 2005) have also been examined. Personal communications with colleagues, internal documents, journal articles, conference proceedings are all sources that engineers use to satisfy their information needs (Hertzum & Pejtersen, 2000). In principle, due to the practical nature of these professions they do not cite and may read less than university faculty members (Tenopir & King, 2000 cited by Tenopir, King, Clarke, Na, & Zhou, 2007). For instance, a survey of paediatricians at the University of Tennessee revealed that they read journal articles for updating their knowledge but read less than do medical faculty members in the same organization (Tenopir et al., 2007). In summary, there is evidence that some professions outside academia read scientific articles and therefore the impact of these articles would not be fully reflected by citations but little is known about how the impact of publications on different professions could be measured.

CHAPTER 3: ASSESSING NON-STANDARD ARTICLE IMPACT WITH F1000 LABELS

3.1 Introduction

The real impact of medical research cannot be captured by citations alone and so new approaches are needed to measure the impact of publications in this area.

Some studies have found significant correlations between FFa scores and citation-related metrics (Bornmann & Leydesdorff, 2012; Li & Thelwall, 2012, Waltman and Costas, 2014). This study examines F1000 from a new perspective: labels assigned to reviewed articles by F1000's reviewers and their relationships with citations and FFa scores and is driven by the research questions below:

1. Does the type of an article impact on its F1000 rating?
2. Does the type of an article impact on its citation count?

3.2 Research Method

F1000 includes two main databases, F1000 biology and F1000 Medicine. In each database, the papers are classified under several sub- disciplines of biology and medicine based on F1000's own classification (see http://f1000.com/prime/recommendations). The research sample is limited to medical articles because the predefined classifications for articles are more suitable for medical research than Biology. The peak time for receiving citations is normally three years after a journal article is published (Moed, 2005). Therefore, in this research, the time span was restricted to articles published in 2007 and 2008. By the time of data collection in January 2012, 3,307 and 5,091, articles published in 2007 and 2008, respectively, had been evaluated by F1000 medicine. In the first step, we extracted the bibliographic information of all 2007 and 2008 articles that had been evaluated by F1000 medicine. In the next step, random samples of 350 out of 3,307 and 550 out of 5,091 records for 2007 and 2008 were selected using the SPSS random sample generator function. This sample covers more than 10% of all reviewed papers from 2007 and 2008 in F1000. Finally, FFa scores and labels of the sampled articles were extracted manually from F1000. Finally, FFa scores and labels of the sampled articles were extracted manually from F1000. The labels available at the time of data gathering were used; however, they are occasionally changed by F1000. At the time of data collection, the F1000 scoring system was different from their current method. Formerly, "the FFa was calculated from the highest rating, which has a value of 10 for Exceptional, 8 for Must Read, and 6 for Recommended. An incremental value was then added for each extra rating (3 for Exceptional, 2 for Must Read, 1 for Recommended)" (Huggett, 2012, p. 8). Figure 3.1 shows the calculation method for the final FFa scores.

Article A									
Rating	Exc	Exc							**FFa**
Score	10	3							**13**
Article B									
Rating	MR	MR	MR	Rec					**FFa**
Score	8	2	2	1					**13**
Article C									
Rating	Rec	Rec	Rec	Rec	Rec	Rec	Rec	Rec	**FFa**
Score	6	1	1	1	1	1	1	1	**13**

Figure 3.1. The F1000 method of calculation for FFa scores (Huggett, 2012)

Figure 3.1 shows the calculation method for three papers. The final FFa score for paper A with two Exceptional ratings is 13 (10 for the first Exceptional score + 3 for the second Exceptional score). The FFa score for paper B is 13 as it was rated as Must Read three times and Recommended once (8 for the first Must Read, 2 for each of the other two Must Read scores, and 1 for the Recommended). The FFa score for paper C is 13 with 8 times Recommended (8 for the first Recommended rate + 1 (×7) for the other Recommended). Waltman criticized the FFa calculation because “incremental recommendations have less weight than the initial recommendation” (Huggett, 2012, p. 9). Based on this suggestion, F1000 have modified the methods rating the reviewed papers and now they rate an article as “'Good', 'Very Good' or 'Exceptional' and assign equivalent to scores of 1, 2 or 3 stars, respectively” (F1000, 2014).

Citation data for each article in the sample was downloaded from Scopus by searching manually using the article title in quotation marks in the article title drop-down menu. Some records were removed as they were duplicated in Scopus with different citation counts, leaving 344 articles from 2007 and 533 from 2008 for this study. Articles were classed as duplicates if they had the same authors, title, publication year and source title. The records removed were only 2.5% of our sample and probably did not affect the results of the analysis. In order to investigate relationships between the main features of articles using FFa scores and citations, the non-parametric Mann-Whitney U-test was used because citations and FFa scores were skewed.

3.3 Results

There is a low but significant correlation between FFa scores and citations for articles published in both 2007 (Spearman r=0.383, n=344, p=0.01) and 2008 (Spearman r=0.300, n=533, p=0.01). It should be noted that r provides information about the strength of the correlations and p gives information about the likelihood of the correlation being non-zero by chance. Here r=0.383 and 0.300 are interpreted as low in terms of the strength of the correlations. In both correlations p=0.01 is an indication of the significance level.

As Table 3.1 shows, most of the sampled articles were evaluated only once by F1000 faculty members. There was a strong correlation (Spearman r=0.509, p=0.01) between evaluation frequencies and FFa scores for all articles studied, which is unsurprising since evaluation frequency contributes to overall FFa scores.

Table 3.1. Frequency of F1000 evaluations for the sampled reviewed articles in 2007 and 2008.

Evaluation frequency	1	2	3	4	6	8
Number of articles	89.5%	8.5%	1.5%	0.3%	0.1%	0.1%

Table 3.2 compares the decisions made by different reviewers for the articles. The reviewers assigned the lowest level of evaluation, "Recommended", to the majority of the papers investigated (65.8%) while only 4.8% were assigned the highest level, "Exceptional".

Table 3.2. F1000 articles based on evaluation times and ratings for articles sampled from 2007 and 2008.

No of Reviews	1	2	3	4	6	8	Total
Articles	784	74	14	3	1	1	877
Exceptional	3% (26)	1% (9)	0.5% (4)	0.1% (1)	0.1% (1)	0.1% (1)	4.8% (42)
Must Read	20% (175)	6.2% (54)	2% 18)	0.3% (3)	0.5% (4)	0.3% (3)	29.4% (257)
Recommended	55.5% (487)	7.6% (66)	1.7% (15)	0.8% (7)	0 (0)	0.4% (3)	65.8% (578)

The majority of 2007 (72%, n = 247) and of 2008 (80%, n = 441) articles were reviewed in the publication year. Approximately 94% (n=822) of the articles were labelled at least once with one of the pre-defined categories. Around 56% (n=450) of the articles received one label, 35% (n=285) were labelled twice, and 11% were labelled three or more times in different categories. It should be noted that a reviewer could assign more than one label to an article. There was a significant but small positive correlation between the number of assigned labels and FFa scores (Spearman r=0.242, p=0.01) as well as between the number of labels and citation counts (Spearman r=0.201, p=0.01).

As shown in Table 3.5, the majority of the tagged articles were labelled “New Finding” (54%) or “Confirmation” (43%), 25% were “Clinical Trial”, 11% were “Controversial”, 10% were “Technical Advance” 8% were “Changes Clinical Practice” and 6% were “Review”. A few articles were labelled “Refutation” (1.5%) or “Novel Drug Target” (0.5%).

In order to control Type I errors, a Bonferroni correction was used for the Mann–Whitney U tests. So, rather than using .05 as the critical level of significance, 0.05/9=0.005 was used (level of significance/number of tests). Exact significances should be used in Mann–Whitney U tests when sample sizes are small (Field, 2009), so in this study, for the “Refutation” and “Novel Drug Target” categories, exact significance is reported.

Descriptive statistics for citation counts and FFa scores for articles 2007 and 2008 are shown in Table 3.3. It is clear that the data for citation counts is skewed.

Table 3.3. Descriptive statistics for citation counts and FFa scores for articles 2007 and 2008

	Citation	FFa scores
N	877	877
Mean	56.46	6.98
Median	27	6
Minimum	0	6
Maximum	2691	20

Table 3.4 provides descriptive statistics for the citation counts and FFa scores of the papers 2007 and 2008 based on their assigned labels. As shown in this table the median citations for those papers classified as “Refutation” was higher than the

median citations for other labels. The median FFa scores for those papers were labelled as Changes to Clinical Practice was higher than for the other categories.

Table 3.4. Descriptive statistics for the citation counts and FFa scores of the papers 2007 and 2008 based on the assigned labels.

Labels	N	Mean (Citation/ FFa)	Median (Citation/ FFa)	Std. Deviation (Citation/ FFa)
New Finding	438	66.72 7.13	34 6	154.42 1.71
Confirmation	351	58.11 7.10	27 6	104.12 1.63
Clinical Trial	206	65.66 7.14	35 6	118.52 1.75
Controversial	93	57.45 7.10	34 6	75.89 2.03
Technical Advance	81	83.67 7.18	24 6	300.83 2.03
Changes to Clinical Practice	62	91.20 8.35	42 8	146.13 2.40
Review	50	69.66 7.26	44 6	93.67 1.50
Refutation	12	73.083 8.33	70 7	43.94 3.93
Novel Drug Target	3	45.66 6.66	38 6	32.19 1.15

The differences between the assigned labels for both FFa scores and citations in most cases were not statistically significant. Table 3.5 shows that both FFa scores and citations of articles labelled “Confirmation”, “Clinical Trial”, “Controversial”, “Technical Advance”, “Review”, “Refutation” and “Novel Drug Target” were not significantly different to those articles which were not categorized as one of these classes (i.e., an indicative p value >0.005). In contrast, there were significant differences in FFa scores for articles which were labelled as “Changes to Clinical Practice” and papers that were not classified in this category. Articles labelled “New Finding” also have significant differences in terms of citation counts with articles that were not classified under this label. The median number of citations of articles classified as “New Finding” (34) was more than for the remaining papers (23). The median FFa scores of articles categorised as “Changes to clinical practice” and other papers was the same (8). The relative merits of significance levels and p values are under debate at the moment, but this thesis adopts the conventional approach with significance levels because it makes it easier to summarise the large numbers of tests reported. Bootstrapping could also have been used but standard statistical tests seemed adequate for the data analysed here.

Table 3.5. Tests for significant differences in citation counts and FFa scores between articles assigned each label and articles not assigned the label. Articles are from 2007 and 2008.

Labels	**FFa Scores**			**Citation Counts**	
	N	Mean Rank	P-value	Mean Rank	P-value
New Finding	438	428.34	0.012	447.04	0.000
Not New Finding	384	392.29		370.97	
Confirmation	351	430.16	0.024	407.61	0.685
Not Confirmation	471	397.59		414.4	
Clinical Trial	206	433.4	0.075	445.08	0.019
Not Clinical Trial	616	404.18		400.27	
Controversial	93	410.45	0.958	437.55	0.261
Not Controversial	729	411.63		408.18	
Technical Advance	81	432.4	0.332	399.15	0.622
Not Technical Advance	741	409.22		412.85	
Changes to Clinical Practice	62	572.92	0.000	480.34	0.018
Not Changes to Clinical Practice	760	398.33		405.88	
Review	50	461.79	0.072	485.56	0.023
Not Review	772	408.24		406.7	
Refutation	12	497.88	0.152	591.75	0.007
Not Refutation	810	410.22		408.83	
Novel Drug Target	3	383.67	0.898	487.17	0.600
Not Novel Drug Target	819	411.6		411.22	

3.4 Limitations

The current investigation has several limitations. Each article could be assigned to more than one discipline in F1000. As a result, the sum of the papers of all the individual sub-disciplines of Medical Sciences (68,627) was almost twice as large as the total number of F1000 medicine articles (35,232). On average, an article was assigned to two sub-disciplines in F1000. Moreover, an article can be appointed to a discipline in F1000 even it has not been reviewed by faculty members from the category. Thus, defining a subject for a reviewed article, particularly for multidisciplinary papers, is difficult. As mentioned above, the sample for this study was selected from the whole of F1000 medicine. Although the average "field citedness" is similar in some clinical medical disciplines (Seglen, 1997), it is not similar for all sub-fields of the medical sciences (Harnad, 1985). Therefore, the citation propensity for the sample investigated from all medical disciplines can affect the results of the present research. Additionally, the final FFa score is calculated based on individual ratings of experts from different disciplines. This means that the papers which were assigned to a discipline in F1000 do not reflect only the activities of the faculty members of that particular category. F1000 reviewers and the number of reviewed papers are not distributed uniformly across disciplines (see Appendix 1). The academic positions of F1000 experts are unknown in this analysis, despite the fact that the academic status of reviewers is an important issue (Zuccala, 2010) in a peer review system. Additionally, F1000 claims that senior scientists and experts review papers but this is not possible to verify. The points scheme used by F1000 to convert judgements to a score is relatively arbitrary and a different scheme would produce a different ranking order for the articles For example, due to the lower weight of incremental additions to the rating compared to the initial rate, the validity of the F1000 scoring system may be illogical (Huggett, 2012, p. 9). Nevertheless, any change in the points scheme would be unlikely to affect the results of this study, because the majority of the articles (89.4%) were only reviewed once, and for these articles the actual scores assigned are irrelevant for their ranking (and as a result for the tests used here). Therefore, a different scoring system would make little difference to the results. Finally, the sample of this study is restricted to papers with F1000 recommendations while including the records without recommendations (zero records) may affect the results, such as correlations. Hence, it is not clear whether it is reasonable to assume that articles not covered by F1000 have a zero score or not.

3.5 Discussion

The significant difference between the citation counts of articles which were classified as "New Finding" and other articles suggests that the 54% of articles with novel findings tend to receive more citations than other articles, presumably because the new findings are more useful for future published research, even though there was no evidence that such articles were more highly rated by the F1000 system. There was a significant difference between the FFa scores of the 8% papers which were classified as "Changes Clinical Practice" and other articles. The appropriateness of medical research for clinical practice is presumably highly valued by experts even though it may not lead to increased citations. It seems that the FFa score is able to recognize appropriate articles for clinical practice better than citations and this is logical because citation practice is restricted to authors' activities while the suitability of an article for clinical use is a separate issue.

Overall, the labels assigned alongside the peer review ratings show that some types of research can be identified by reviewers as being, on average, more valuable than others even if this will not be recognised through citations. Therefore, F1000 could be used in research evaluation exercises when the importance of practical findings needs to be recognised. Furthermore, since the majority of the articles studied were reviewed in their publication year, F1000 could be useful when quick evaluations are needed. However, in agreement with Waltman and Costas (2014), the idea of using F1000 as a source for research evaluation needs a deeper analysis because of challenges and unknown issues of F1000.

CHAPTER 4: MENDELEY READERSHIP FOR RESEARCH EVALUATION AND KNOWLEDGE FLOWS

4.1 Introduction

Although previous studies have found significant medium correlations between citations and Mendeley readership counts for specific sets of articles, it seems that no previous research has investigated the relationship between Mendeley readership counts and citation measures across a range of disciplines. This is an important omission because citation behaviours across disciplines are known to vary and so Mendeley readership counts may not always correlate with citation counts. This chapter partly fills this gap by investigating the correlation between Mendeley readership and citation counts for different social sciences and humanities disciplines. Mendeley readership data is also used to discover relationships between social sciences and humanities disciplines (i.e., on a much larger scale than previously attempted with Mendeley), assessing whether the results are reasonable through comparisons with cross-disciplinary citations. The following research questions drive the investigation:

1. Are there significant, substantial and positive correlations between Mendeley readership counts and citation measures in all social sciences and humanities disciplines? If so, are there significant differences between disciplines?
2. To what extent do Mendeley bookmark data reflect similar information flow patterns to the cross-disciplinary citations in WoS?

4.2 Research Method

The advance search options of WoS online were used to retrieve all social sciences and humanities articles indexed by WoS in the year 2008. The results were limited to research articles in English (removing reports, editorials, book reviews, etc.). The year 2008 allows enough time for citations to accrue because the peak time for citations is usually three years after an article is released (Moed, 2005). However, some disciplines in the humanities are slower to reach the peak citation rate. In order to classify the records into social sciences and humanities disciplines, the "Analyze Results" functionality was used, then refining the documents to "Research Areas", which is abbreviated to "SU" in the WoS interface[9]. Finally, ten social sciences and humanities disciplines were selected to give a wide range, as shown in Table 4.1. In other words, scientific disciplines were operationalized based on WoS Research Areas. The level of aggregation for WoS Research Areas is higher than for WoS categories (Leydesdorff, Carley, & Rafols, 2013). For instance, all sub-disciplines of Psychology are classified as a single discipline named Psychology. In the next step, bibliographic information and citation counts for all publications in the selected disciplines were downloaded from WoS for further analysis.

Mendeley users save reference information in their libraries. For the purposes of this chapter, the number of Mendeley users that had saved a paper is described as that paper's *Mendeley readership count*, whether they had read the paper or not. In order to obtain the readership data, Webometric Analyst [10] automatically extracted Mendeley readership counts for the WoS articles via the Mendeley API. As multiple versions of an article sometimes exist in Mendeley, duplicate records were identified

[9] http://images.webofknowledge.com/WOKRS510B3_1/help/WOS/hp_research_areas_easca.html
[10] lexiurl.wlv.ac.uk

based on Mendeley unique IDs, Mendeley URLs, and DOIs. Possible duplicates were manually checked and removed. Out of 41,624 Mendeley records, 1,166 (3%) were judged to be duplicates. Some of the articles in the Mendeley catalogue did not have readership statistics and instead of statistical data the phrase "Readership statistics are being calculated" was displayed. It appears that Mendeley loaded these articles straight from the publishers' websites or that some users had added their own publications to their Mendeley profiles but no one had saved these articles into a personal library, or perhaps these articles had previously been in a personal library. Most of the records removed due to duplication did not have readership statistics.

Table 4.1 shows that an average of 44% of the articles from the chosen social sciences disciplines appeared in the Mendeley catalogue, in comparison to only 13% of the humanities articles. Psychology (54%) and Linguistics (34%) had the highest coverage in Mendeley within the social sciences and humanities, respectively. Education (39%) and Literature (4%) had the lowest percentages of articles in the Mendeley database. In total, 27,558 social science articles and 1,914 humanities articles with Mendeley readership statistics were used. Spearman correlation tests were applied to the WoS citations and Mendeley readership counts. Spearman correlation was preferred to Pearson correlation because the frequency distributions of readership and citation counts are skewed.

This chapter also uses citations and Mendeley readership counts for publications to measure information flow across different disciplines. For this purpose, the same search methods for retrieving the data from SSCI and AHCI were used and the "create citation report" option was used to extract citing disciplines based on WoS Research Areas for the selected social science and humanities disciplines. To compute information flows based on Mendeley readership, users' research backgrounds in Mendeley profiles were used. Data are accessible through the Mendeley API for only the three most common readers' background disciplines for each individual article, however. The data is provided as percentages rather than total numbers of readers. For each article and each of the three readers' disciplines returned, the percentage of readers from the discipline was multiplied by the total number of readers of the article and divided by 100 to obtain the estimated number of article readers from that discipline. This process covered 89% and 82% of the readers' background disciplines for social science and humanities articles, respectively.

Table 4.1. Coverage of WoS articles from social sciences and humanities disciplines in Mendeley.

WoS discipline	*Articles indexed by WoS in 2008*	*WoS articles covered by Mendeley (%)*	*Articles with readership statistics (%)*	*Articles without readership statistics (%)*
Psychology	23,811	14,757 (62%)	12,804 (54%)	1,953 (8%)
Social Sciences Other Subjects	6,366	3,763 (59%)	2,416 (38%)	1,347(21%)
Education and Educational Research	7,208	3,839 (53%)	2,796 (39%)	1,043(14%)
Information Science and Library Science	2,552	1,617 (63%)	1,343(53%)	274 (10%)
Business and Economics	22,710	12,337 (54%)	8,199 (36%)	4,138 (18%)
Social sciences total	***62,647***	***36,313 (58%)***	***27,558 (44%)***	***8,755(14%)***
Philosophy	2,833	1,060 (37%)	468 (17%)	592 (21%)
History	2,882	756 (26%)	253 (9%)	503 (17%)
Linguistics	2,245	1,046 (47%)	773 (34%)	273 (12%)
Literature	4,622	643 (14%)	165 (4%)	478 (10%)
Religion	2,058	640 (31%)	255(12%)	385 (19%)
Humanities total	***14,640***	***4,145 (28%)***	***1,914 (13%)***	***2,231 (15%)***

4.3 Results

The data was analysed separately for the two research questions.

4.3.1 Correlations between WoS Citations and Counts of Mendeley readers

We measured correlations between Mendeley readership counts and citations for all articles with Mendeley readership and articles not covered by Mendeley separately because of the restriction of the analysis to records with Mendeley readership might be misleading if articles that were not in Mendeley tended to be substantially cited and so we performed the analysis with including articles not covered by Mendeley. In the first analysis, the readership counts of those articles that were not included in Mendeley were set at zero. As shown in Table 4.2, there are low and significant correlations for both social sciences ($r=0.350$, $p<.01$) and humanities ($r=0.280$, $p<.01$) disciplines overall. Here, r=0.1+, 0.3+, 0.5+ are interpreted as small, medium, and large correlations, respectively (Cohen, 1988), and medium and large correlations are considered to be substantial. Information Science and Library Science (IS&LS) ($r=0.369$, $p<.01$) and Linguistics ($r=0.162$, $p<.01$) had the highest correlations, while social sciences Other Subjects ($r=0.293$, $p<.01$) and Literature ($r=0.162$, $p<.01$) had the lowest, among social sciences and humanities disciplines respectively.

Table 4.2. Spearman correlations between citations and Mendeley bookmarking counts for all WoS articles from 2008 in different social sciences and humanities disciplines.

Selected WoS discipline	*WoS citation median, mean, Standard deviation*	*Mendeley readership median, Mean, Standard deviation*	*Correlation between Mendeley readership and citation counts*
Psychology	5 8.41 14.19	1 5.25 9.81	.311**
Social Sciences Other Subjects	2 4.20 6.11	0 2.45 5.88	.293**
Education and Educational Research	2 4.0 6.11	0 2.45 5.88	.308**
Library and Information Science	2 5.10 7.80	1 5.88 10.68	.369**
Business and Economics	3 5.50 8.84	0 3.78 9.14	.359**
All selected social sciences	*3* *6.32* *10.89*	*0* *4.22* *9.03*	*.350***
Philosophy	0 1.15 2.33	0 0.78 2.50	.215**
History	0 0.78 1.60	0 0.34 2.06	.193**
Linguistics	1 1.95 3.32	0 1.76 3.88	.312**
Literature	0 0.35 1.19	0 0.15 1.45	.162**
Religion	0 0.70 1.64	0 0.42 1.54	.222**
All selected humanities	*0* *0.89* *2.09*	*0* *0.60* *2.38*	*.280***

** Significant at p = 0.01

In the second analysis, correlations between Mendeley readership counts and citations were limited to those articles with Mendeley readership counts. Table 4.3 shows that there is a significant correlation between Mendeley readership and citation counts in all of the disciplines investigated. The correlation for social sciences disciplines overall (r=0.516, p<.01) is higher than that for humanities disciplines (r=0.428, p<.01). The correlations for social sciences disciplines were medium, varying from r=0.403, p<.01 (Social Sciences Other Subjects) to r=0.573, p<.01 (Business and Economics). Amongst the humanities disciplines, Religion and Philosophy have the lowest correlations (r=0.363, p<.01 and r=0.366, p<.01) and Linguistics has the highest correlation (r=0.454, p<.01).

Table 4.3. Correlations between WoS citations and Mendeley readership counts (non-zero only) for 2008 articles.

Selected WoS discipline	*WoS citation median, mean, Standard deviation*	*Mendeley readership median, Mean, Standard*	*Correlation between Mendeley readership and citation counts*
Psychology	6 9.59 17.21	6 9.69 11.60	.514**
Social Sciences Other Subjects	4 5.62 7.28	4 6.48 8.08	.403**
Education and Educational Research	4 5.71 7.54	6 8.19 9.05	.484**
Information Science and Library Science	4 6.20 8.56	8 11.30 12.56	.535**
Business and Economics	5 7.92 10.73	7 10.56 12.70	.573**
All selected social sciences	5 8.18 13.72	6 9.59 11.56	*.516***
Philosophy	1 2.10 3.34	4 4.77 4.36	.366**
History	1 1.96 3.29	2 3.94 5.86	.428**
Linguistics	2 3.01 4.37	4 5.16 5.15	.454**
Literature	0 1.56 3.67	2 4.46 6.36	.403**
Religion	1 1.56 2.56	3 3.42 3.01	.363**
All selected humanities	1 2.33 3.77	3 4.61 4.99	.428**

** Significant at p = 0.01

Comparing Table 4.2 and Table 4.2 indicates that the correlations between Mendeley readership counts and citation counts are weaker for both social sciences and humanities when we considering records with zero readership counts. The comparison between the two tables also reveals that the median number of citations for articles with zero Mendeley bookmarks for all social sciences disciplines is 2 while the median for the articles with readership counts in Mendeley is 5. For humanities disciplines, the median number of citations to articles with no Mendeley readers is 0, while the median for articles with readership counts in Mendeley is 1. The small difference between these two medians may be the reason for the decreased correlations between citations and Mendeley bookmarks in the second analysis.

4.3.2 Knowledge Flows Based on Mendeley Readers

The remaining tables address the second research question and compare article disciplines with those of their readers. The WoS subject categories do not match the Mendeley categories[11] and so the comparisons are approximate, however. To aid the analysis, similar Mendeley categories were merged into broader categories, and categories with low readership percentages were merged into a single 'Others' category. Detailed information about the merged Mendeley categories is provided in appendices 1 and 2. For citing disciplines, as shown in tables 4.5 and 4.7, small citing research areas have been merged into the main disciplines (see in Appendices 3 and 4 for more details). For example, all sub-fields of medicine were merged into a broader Medicine and Health category. From Table 4.4, we can see the majority of readers of all five selected WoS social science disciplines are from the home

[11] http://www.mendeley.com/directory/

disciplines, except for IS&LS and the broad Social Sciences Other Subjects category. However, the percentages vary across different disciplines, from Psychology (64%) to Social Sciences Other Subjects (28%). This suggests that most Mendeley readers use scientific information mainly from their own disciplines, but that this varies substantially across disciplines. Many readers of IS&LS articles (46%) were recorded as computer and information scientists but these might be more commonly computer scientists than library scientists.

Very few Psychology articles have an Arts and Humanities readership while some Psychology literature is read by people from Biology (7%) and Medicine (6%), perhaps reflecting crossover fields such as neuropsychology and psychopharmacology.

Table 4.4. Mendeley *readership* categories for year 2008 articles from the five selected social sciences. Home categories are in bold for each WoS discipline; values above 10% are shaded.

		Selected WoS discipline				
		Psychology	*Social Sciences Other Subjects*	*Education and Educational Research*	*IS&LS*	*Business and Economics*
Mendeley reader category	Psychology	**64.0%**	15.8%	12.4%	1.8%	6.5%
	Social Sciences	6.5%	**27.8%**	7.4%	20.5%	11.6%
	Education	3.8%	5.4%	**54.4%**	4.4%	1.0%
	Computer and Information Science	3.1%	4.5%	9.0%	**45.9%**	4.7%
	Business and Economics	3.5%	11.6%	1.9%	14.0%	**55.7%**
	Management	0.9%	3.1%	0.5%	3.5%	11.0%
	Medicine	6.1%	7.7%	4.9%	3.1%	1.0%
	Biological Sciences	6.6%	4.5%	1.7%	1.4%	1.5%
	Philosophy	0.4%	4.5%	0.2%	0.1%	0.1%
	Linguistics	1.9%	0.1%	3.0%	0.2%	0.0%
	Art and Humanities	0.5%	2.7%	1.0%	1.4%	0.2%
	Others	2.6%	12.3%	3.4%	3.6%	6.6%
	Total	*112,898*	*13,436*	*20,817*	*13,000*	*74,080*

Table 4.5 indicates the disciplines that cited social sciences articles published in 2008. For most of these disciplines, the majority of papers were cited by publications in the home disciplines, except for IS&LS. The percentage of IS&LS papers cited by IS&LS publications is 24%, which is less than the percentage of Computer Science documents (25%) that cited IS&LS papers.

Table 4.5. Disciplines *citing* articles social sciences articles published in 2008. Home disciplines are in bold for each WoS discipline; values above 10% are shaded.

		Selected WoS discipline				
		Psychology	*Social Sciences Other Subjects*	*Education and Educational Research*	*IS&LS*	*Business and Economics*
Citing discipline in WoS	Psychology	**33.4%**	8.7%	12.5%	2.6%	4.3%
	Social Sciences	7.7%	**32.0%**	7.5%	6.0%	11.7%
	Education	2.9%	2.3%	**36.2%**	2.5%	0.8%
	Information Science and Library Science	0.2%	0.6%	0.8%	**24.0%**	1.3%
	Business and Economics	2.8%	8.6%	2.1%	9.9%	**43.1%**
	Management	0.2%	1.6%	0.5%	3.0%	7.0%
	Medicine and Health	39.1%	25.6%	24.9%	11.7%	5.8%
	Biology and Life Sciences	7.0%	5.5%	2.7%	2.8%	5.8%
	Computer Science	1.1%	1.6%	4.5%	25.0%	3.9%
	Engineering	0.9%	3.1%	2.0%	4.9%	4.7%
	Linguistics	1.2%	0.2%	2.7%	0.1%	0.0%
	Art and Humanities	0.6%	3.8%	1.1%	0.6%	0.5%
	Others	2.7%	6.2%	2.5%	6.8%	11.0%
	Total	*272,957*	*39,926*	*36,520*	*16,751*	*167,996*

Table 4.6 shows that the main users of Philosophy (32%) and Linguistics (55%) publications are from the same discipline. The situation of Literature is unusual because 27% of readers are in the broad Arts and Literature category but 28% are in the Humanities category that is also relevant to Literature. In contrast, the main users

of History (40%) and Religion (27%) articles are from social sciences rather than their home disciplines, which is within the humanities in both cases.

Table 4.6. Mendeley *readership* categories for year 2008 articles from the five chosen Humanities disciplines. Home categories are highlighted for each WoS discipline.

		Selected WoS discipline				
		Philosophy	*History**	*Linguistics*	*Literature*	*Religion**
Mendeley reader category	Philosophy	**32.1%**	4.0%	1.2%	0.9%	6.6%
	Humanities	7.2%	**31.7%**	4.7%	27.8%	**23.1%**
	Linguistics	2.6%	0.7%	**55.0%**	1.2%	2.5%
	Arts and Literature	2.6%	3.8%	2.5%	**27.3%**	1.7%
	Social Sciences	12.4%	39.6%	7.8%	20.6%	26.9%
	Psychology	15.6%	6.5%	8.4%	1.3%	21.4%
	Education	3.7%	2.4%	7.9%	2.6%	6.4%
	Business and Economics	1.1%	1.2%	0.1%	1.0%	1.1%
	Medicine	2.4%	0.7%	0.5%	1.0%	3.4%
	Biological Sciences	5.0%	0.7%	0.9%	0.6%	2.3%
	Computer and Information Science	6.5%	2.8%	9.3%	10.1%	1.1%
	Others	8.8%	5.9%	1.7%	5.6%	3.5%
	Total	*1153*	*911*	*3760*	*650*	*812*

*History and Religion are sub-categories of the Mendeley Humanities category (http://www.mendeley.com/disciplines/humanities/).

Comparing information flows based on citing and Mendeley bookmarking for humanities disciplines shows that they are similar in most cases. Nevertheless, there are some differences in terms of the strengths of connections between the disciplines. For instance, Mendeley usage data gives stronger connections between History, Psychology and Computer Science while citation data shows stronger relationships between History, Medicine and Biology. However, this may be related to the time lag in receiving citations in these disciplines.

Table 4.7. Disciplines *citing* humanities articles published in 2008. Home disciplines are highlighted for each WoS discipline.

		Selected WoS discipline				
		Philosophy	*History*	*Linguistics*	*Literature*	*Religion*
Citing discipline in WoS	Philosophy	**46.4%**	3.1%	1.1%	2.7%	4.00%
	History	0.9%	**35.0%**	0.5%	3.6%	4.3%
	Linguistics	2.3%	0.6%	**48.3%**	8.3%	0.2%
	Literature	1.1%	2.0%	3.0%	**44.5%**	1.0%
	Religion	1.6%	2.5%	0.1%	1.2%	**39.0%**
	Arts and Humanities Other Subjects	1.1%	4.7%	1.1%	10.7%	1.8%
	Social Sciences	17.6%	34.5%	9.8%	17.5%	23.9%
	Psychology	6.2%	1.2%	11.0%	1.6%	8.7%
	Education	1.7%	0.9%	7.6%	2.8%	4.2%
	Business Economics	1.5%	6.8%	0.4%	0.7%	0.7%
	Medicine and Health	4.0%	1.9%	6.2%	1.0%	7.0%
	Biology and Life Sciences	3.4%	2.4%	1.9%	1.2%	1.7%
	Computer Science	3.6%	0.5%	4.3%	1.4%	1.2%
	Others	8.5%	3.8%	4.70%	2.7%	2.5%
	Total	*4,222*	*2,829*	*5,282*	*1,876*	*1,822*

4.4 **Limitations**

A limitation of this study is that readership is limited to the individuals who choose Mendeley as their reference manager; many scholars use other similar tools, such as EndNote, RefWorks and ProCite, to organize their references, or do not use a reference manager at all. Another limitation is that around 11-18% of the readers' background disciplines were not accessible via the Mendeley API, and so the contributions of minor subjects and the extent of interdisciplinarity may be under-estimated. The results also reflect the size of the disciplines involved and the extent to which Mendeley is used within these disciplines. Hence, the results are likely to be skewed towards disciplines using Mendeley the most actively (e.g., perhaps IS&LS).

The WoS research areas used to define disciplines and the overlap between WoS subjects are also a limitation of this research, because 25% of WoS journals have more than one subject classification (Rinia, Van Leeuwen, Bruins, Van Vuren, & Van Raan, 2002). However, it is not easy to label research publications as belonging to a single subject, particularly for multidisciplinary research areas and disciplines with common research boarders. The classification of citation flows in this chapter is based on WoS research areas, which operate at the journal level, while the categorization for cross-readership is based on the research interests of Mendeley users. As a paper could appear in more than one WoS research area category but Mendeley users can choose only one research interest for a paper, this issue can affect information flows in terms of readership and citation. For example, a paper that has been classified in two disciplines in WoS but only one in Mendeley will appear to be part of an information flow between areas for at least one of the WoS disciplines.

Finally, the sample is restricted to journal articles only, although books are a fundamental source of research in many humanities and some social science disciplines (Huang & Chang, 2008; Nederhof, 2006). However, social sciences and humanities researchers have begun to publish more in WoS journals (Kyvik, 2003; Butler, 2003)

4.5 Discussion

This chapter examined Mendeley usage data for social sciences and humanities publications from 2008. In answer to the first research question, there were statistically significant medium positive correlations between Mendeley readership counts and citations for all the selected disciplines but the values varied across disciplines. The highest correlations between Mendeley readership counts and citations are in those disciplines that are closer to hard sciences in terms of citation behaviour, while correlations are lower in the disciplines which more closely resemble traditional humanities. The median Mendeley readership counts were higher than the median citations for the articles covered by Mendeley in all the disciplines except Psychology. In almost all disciplines, the correlation is not strong enough to conclude that Mendeley readership and citation counts measure the same aspect of research impact. As hypothesised by other researchers, a likely explanation is that Mendeley captures broader scholarly activities from a variety of readers' perspectives in comparison with citation counts.

In answer to the second question, comparing knowledge transfer results across the disciplines examined based on citation data with Mendeley bookmark data shows that both datasets have similar overall patterns for the disciplines investigated. Nevertheless, there were some differences in the strength of links between scientific subjects. For instance, Mendeley bookmark data gives stronger interdisciplinary connections in the humanities and most of the social sciences. Thus, the results of this study support the value of using Mendeley readership data to discover meaningful knowledge transfer patterns across scientific disciplines. This confirmation is evidence that Mendeley is a reasonable tool to measure information flow across scientific disciplines. This finding is in agreement with previous studies (Jiang, He, & Ni, 2011; Kraker, Körner, Jack, & Granitzer, 2012) which showed that social bookmarking tools provide a valuable source for discovering relationships between disciplines but the current study examined broader scientific disciplines.

CHAPTER 5: MENDELEY READERSHIP CATEGORIES

5.1 Introduction

In order to holistically evaluate the use of research, it is important to know who reads academic articles and why. These issues have not been systematically examined before because of anonymity in usage data for electronic journals and because of the lack of usage data for print journals. Although some studies have explored article readership on a small scale (Niu & Hemminger, 2012; Hemminger, Lu, Vaughan, & Adams, 2007) there are no large-scale systematic investigations into what types of people read scholarly articles. However, some studies have used Mendeley users' occupations for altmetric investigations with samples of articles with DOIs (Zahedi, Costas and Wouters, 2013) and articles in the Journal of Strategic Information Systems (Schloegl, Gorraiz, Gumpendorfer, Jack, & Kraker, 2013). The first of the above references (a workshop presentation rather than a published paper, and so full details of the methods and results are not available to be analysed and it is not fully peer-reviewed) is the closest to the current chapter. It includes an analysis of a random sample of 200,000 WoS articles from 2011-2012, revealing the typical occupations of article readers, the types of document that the different occupations read, a broad disciplinary breakdown of the users into seven categories, a comparison of the number of readers and citations for each broad category, and a comparison between the number of readers and the number of citations for each occupation. The current chapter reports similar findings but at a more fine-grained level, analysing individual disciplines rather than broad groups of disciplines, with a larger sample size, and also reporting additional types of analyses. It analyses the occupations of readers (e.g., professors, PhD students, undergraduates, non-academic users) of research articles for several disciplines in Mendeley. Additionally, the effects of users' occupations on correlations between Mendeley readership counts and citations are investigated. The following research questions drive this chapter, focusing on several broad areas of science.

1. What are the common types of readers for Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry research articles in Mendeley?
2. Does the academic or professional status of readers in Mendeley affect the relationship between Mendeley readership counts and citation counts?

5.2 Research Method

The *Observatoire des Sciences et des Technologies'* in-house version of the Thomson Reuters WoS databases was used for lists of articles in academic journals. All bibliographic information and citation data for WoS journal articles from 2008 were selected, excluding non-article document types, such as editorials and book reviews. The citation data comes from the *Science Citation Index Expanded (SCIE)*, the *Arts and Humanities Citation Index (AHCI)* and the *Social Science Citation Index (SSCI)* in December 2012. Both of these are widely used in scientometric research and are accepted to have good coverage of high impact academic journals, although with some biases, such as towards English language journals. The choice of the year 2008 allowed all articles at least four years to receive citations. For defining the main research disciplines and sub-disciplines, the US National Science Foundation (NSF) classification was used. This classification is more suitable for this study than the

WoS classifications because each journal is assigned to only one NSF research speciality or sub-discipline. The 22 most productive disciplines in terms of the number of publications in the year 2008 were selected for the study from the broad NSF categories of Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry (see Appendix 2). These disciplines include 44% of the journal articles from 2008 in the Thomson Reuters databases used.

Next, using the Mendeley API, Mendeley readership counts for each selected WoS article were automatically extracted with Webometric Analyst[12] using a query consisting of the last name of the first author, publication year and title of the article. Instead of relying on a document identifier such as a DOI, which is often missing in the Mendeley entries, this method increases recall by relying on three main metadata elements. However, documents with at least one incorrect item of bibliographic information (e.g., author or year) were ignored to increase precision. As multiple copies of a paper could exist in Mendeley, duplicate records were identified and removed based on WoS unique IDs. Out of 480,979 WoS articles for all disciplines, 219,326 (45.6%) were found in Mendeley and 3,745 were duplicates. Removing duplicates reduced the overall readership count by 1.1% (see Appendix 6). This process can cover documents without DOIs but is a time-consuming way to match data sets.

Mendeley coverage varied by discipline. Clinical Medicine articles had the highest coverage in Mendeley with 71.6% having a Mendeley record (see Table 5.1), while in Physics, Chemistry and Engineering and Technology only about one third of the documents were saved in Mendeley. Mendeley records with zero readers in the database were disregarded. These papers could have been added to Mendeley in

[12] lexiurl.wlv.ac.uk

several ways. For instance, Mendeley may automatically add all articles from specific publishers. Moreover, some journals administrators or researchers may add all their publications to Mendeley to increase their visibility (for example, see, http://www.mendeley.com/profiles/ijcsi-journal). Detailed information for articles with zero readers is listed in Table 5.1. As shown in Table 5.1, 41.1% of the WoS articles had Mendeley readership statistics. All of the 197,848 of the WoS articles with Mendeley readership statistics from Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry were selected for further analysis (see Appendix 6).

Table 5.1. Coverage of WoS articles from Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry in Mendeley.

Discipline	Articles indexed by WoS in 2008	Unique WoS articles covered by Mendeley	Duplicate WoS records in Mendeley	Articles with readership statistics in Mendeley	Articles without readership statistics
Clinical Medicine	145,536	104,189 (71.6%)	1,618 (1.5%)	90,358 (62.1%)	13,831 (9.5%)
Engineering and Technology	109,390	38,082 (34.8%)	582 (1.5%)	35,633 (32.6%)	2,449 (2.2%)
Social Science	23,878	11,172 (46.8%)	567 (4.8%)	10,948 (45.9%)	224 (0.9%)
Physics	101,581	31,921 (31.4%)	393 (1.2%)	30,124 (29.7%)	1,797 (1.8%)
Chemistry	100,594	33,875 (33.7%)	585 (1.7%)	30,785 (30.6%)	3,090 (3.1%)
Total	480,979	219,239 (45.6%)	3,745 (1.7%)	197,848 (41.1%)	21,391 (4.4%)

Although the Mendeley API provides information related to the discipline, academic status and country of readers for each record, it only reports percentages rather than raw data and only gives information about the top three categories. For each article and each of the top three readers' occupations for that article, the percentage of readers with that occupation was multiplied by the total number of readers of the article and divided by 100 to obtain the estimated number of article readers from that occupation.

To understand the extent to which the three most frequent statuses represented the entire readership of a document, the sum of the (up to) three status percentages was subtracted from the total readership counts to indicate the missing information per document. As shown in Table 5.2, academic status information was not available for 27% of the readers due to the restrictions of the Mendeley API.

Table 5. 2. Available and missing Mendeley user status information for readership counts for Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry via the Mendeley API.

Discipline	Total readership counts	Readership counts with status information available via the API for the top 3 categories	Readership counts without status information
Clinical Medicine	699,681	70.5%	29.4%
Engineering and Technology	324,624	75.2%	24.7%
Social Science	140,952	69.0%	31.1%
Physics	251,071	76.5%	23.4%
Chemistry	231,313	76.9%	24.3%
Total	1,647,641	73.1%	27.0%

Some of the 15 occupational categories reported by Mendeley are similar and were merged into a single category. For instance, postgraduate students and masters students were merged into a single postgraduate student category (see Appendix 7).

5.3 Results

5.3.1 Readers and Occupations

Because Mendeley only reports reader counts for the top 3 occupational categories for each article and this biases the results so that they underestimate the percentages of categories which frequently do not belong to the top 3, results are also provided for documents where the top 3 categories made up 100% and at least 66% of all reader counts, respectively. As the actual unbiased percentage of readers per status cannot be exactly determined based on the data provided, the three values can thus be considered as estimates of the actual values, where the true figure lies somewhere between the three values for each occupation.

Figure 5.1 shows that in all disciplines PhD students were the main Mendeley readers of articles in 2008 for all papers, papers with at least 66% and papers with 100% reader counts, although the percentages vary by discipline. Postgraduate students and Postdocs were the main readers after PhD students across the different disciplines, as shown in Figure 5.1 and appendices 7, 8 and 9. All of the professions are self-reported and it is possible that, for example, some of the people recorded as *Professor* might not be full professors. Moreover, people with other academic ranks, such as Reader or Lecturer in the UK, might not map themselves accurately to the most similar Mendeley category. Finally, it is possible that some users may have been promoted since registering their status (e.g. PhDs to Postdocs) without changing their professions in their Mendeley profiles.

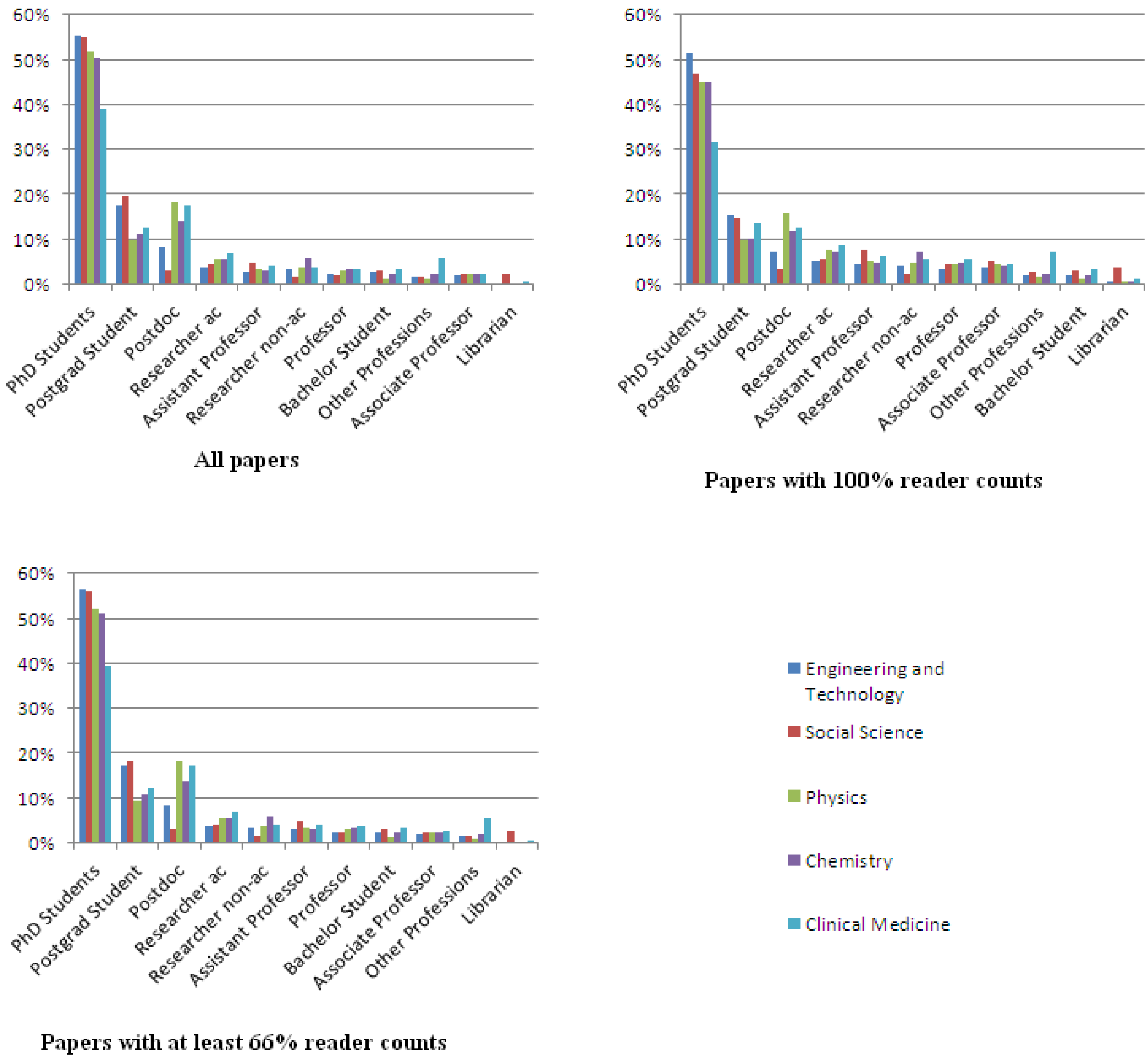

Figure 5.1. Readers of WoS articles from Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry in Mendeley based on academic status for all papers, and for papers with 66% and 100% readership counts.

Among the selected disciplines, about 7.2%, 5.9% and 5.6% of the readers of Clinical Medicine papers were from the Other Professions category for papers with 100%, all papers and papers with 66% reader counts respectively. Librarians were 3.7%, 2.8% and 2.5 % the reported readers of Social Sciences articles but were the least common readers of papers in other disciplines.

5.3.2 Correlations between Mendeley Readership Counts and Citations based on Users' Occupations

Spearman correlations were calculated between Mendeley readership counts and citations for all articles with at least one reader in Mendeley for each sub-discipline. As in the previous chapter, values of r = 0.1+, 0.3+, and 0.5+ (whether positive or negative) were considered to be small, medium, and large, respectively (Cohen, 1988), with medium and large correlations considered to be substantial. There were statistically significant positive correlations between Mendeley readership counts and citations for all five disciplines (Table 5.3, see also Appendix 11). The correlations for all of the disciplines are similar but are highest in Clinical Medicine and Social Science (r=.561).

In order to investigate the effect of non-read articles on correlations between Mendeley readership counts and citations, the analysis was repeated but including articles with zero readers (including articles that were in Mendeley but which were not found by the search process). In other words, all articles not found in Mendeley were assumed to have zero Mendeley readers. As shown in Table 5.3 the correlations between Mendeley readership counts and citations are weaker for all disciplines.

The median Mendeley readership counts for all Social Science sub-disciplines are higher than their median citation counts, and the overall median Social Science Mendeley readership count is double the median Social Science citation count (Table 5.3). The opposite is true for Physics, Chemistry and Clinical Medicine but Engineering and Technology follow the pattern of the Social Sciences. If the dataset had included articles with zero Mendeley readership counts, then the correlations would probably have been weaker, as confirmed below and in Table 4.2.

Table 5.3. Spearman correlations between WoS citations and Mendeley readership counts (both zero and non-zero) for 2008 articles.

Main disciplines	WoS citation median non-zero only/ Both zero and non-zero	Mendeley readership median non-zero/ Both zero and non-zero	Spearman correlation non-zero/ Both zero and non-zero
Clinical Medicine	9 7	4 2	.561** .463**
Engineering and Technology	5 3	5 0	.501** .327**
Social Science	4 2	8 0	.561** .456**
Physics	7 4	5 0	.548** .308**
Chemistry	11 6	5 0	.554** .369**

**Significant at the level of p=0.01

As mentioned above, only the top three occupations for Mendeley readership counts are available for each article. To try to overcome this limitation, correlations were calculated between Mendeley readership counts and citations for several professions for three datasets based on the availability of readership data: a) all articles, b) articles with at least 66% of reader occupations available, and c) articles with 100% of reader occupations available. The correlations for all papers are presumably overestimates, especially for those occupational categories that often do not belong to the top 3 reported ones. As the actual unbiased correlation values cannot be computed, the three values are considered to be estimates, where the 100% value reflects the lower bound of the correlations.

There are positive correlations between Mendeley readership counts and citations for all occupations except librarians for some sub-disciplines, although the strengths of the correlations vary by occupation across research disciplines (Figure 5.2). As shown in Figure 5.2, the correlations decrease for records with 66% of the readership occupations available in comparison to all articles and the correlations are smaller for records with 100% of the readership occupations available in contrast to the sets of articles with at least 66% of readership occupations available. In other words, all correlations are lower for papers with 100% of readership occupations available. The likely reason for this is that these are the least cited papers, with the lowest total number of readers, and so the correlation test is less powerful for them because the numbers are smaller. Generally, the highest correlations are for full professors, assistant professors, postdocs and PhD students, while the lowest correlations are for undergraduates, other professions and librarians in all disciplines in all three datasets. The pattern of correlations for researchers at academic and non-academic institutions is similar across the research areas for all three datasets. However, the differences between correlations for undergraduate and postgraduate students are noticeable for all disciplines (see appendices, 11, 12 and 13). The correlations between Mendeley readership counts and citations for full professors, assistant professors, post docs, PhD students and postgraduate students are substantial for all disciplines. As shown in Figure5.2, the correlations for undergraduates and other professions are small. Nevertheless, the correlations for *other professions* are higher for Clinical Medicine than the other disciplines.

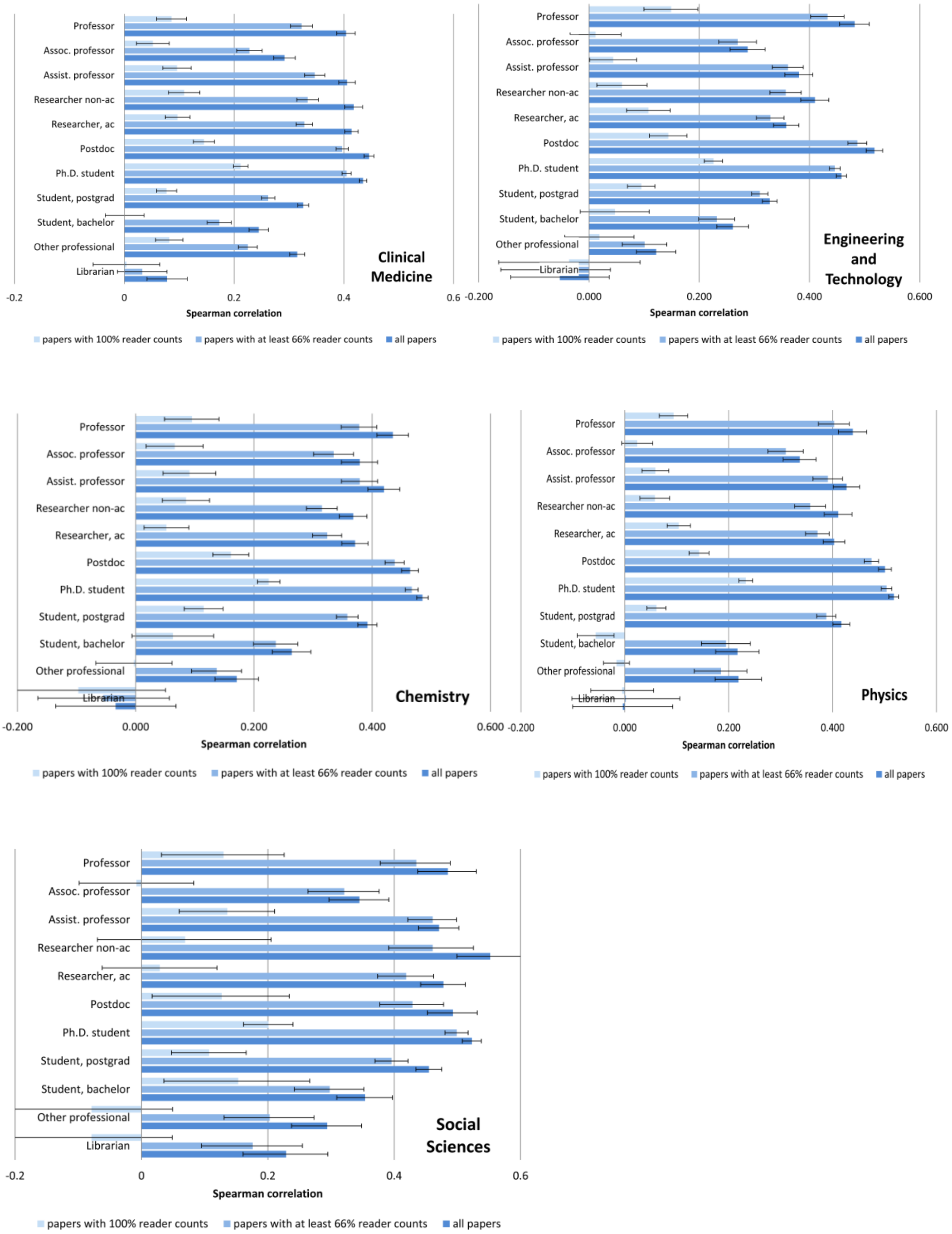

Figure 5.2. Spearman correlations between Mendeley readership counts and citations based on occupations for Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry. Figures are reported separately for three datasets, determined by the percentage of reader occupations known for an article. Error bars give a 95% confidence interval, calculated using a Fisher transformation of the correlation to give it an approximately normal distribution.

5.4 Discussion

Most readers of Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry papers in Mendeley are PhD students, which is in agreement with previous findings (Zahedi, Costas and Wouters, 2013). Postgraduate students and postdoctoral researchers are the two most common readers of papers in Mendeley across different disciplines, after PhD students. Perhaps the most important reason for this is that Mendeley attracts young researchers because they have more energy and time to start a new system than do older scholars. Another possible explanation is that PhD students and postdoctoral researchers mainly focus on research, whereas the other groups are likely to have additional responsibilities. In addition, PhD students use more references in their publications than do faculty members (Larivière, Sugimoto, & Bergeron, 2013). PhD students and postdoctoral researchers also mine related literature more than do senior researchers, as they try to obtain comprehensive knowledge about their research topics while older researchers are often co-authors (Gingras, Larivière, Macaluso, & Robitaille, 2008) and may have supervisory roles in research projects. Alternatively, younger researchers could be more adaptable to novel ideas and read more new publications, while senior scholars use older literature (Barnett & Fink, 2008). Moreover, Mendeley is a new tool and senior researchers seem to avoid using most social web services (Mas-Bleda, Thelwall, Kousha, & Aguillo, in press) preferring to continue with their existing referencing practices.

Postgraduates are readers of many articles in almost all of the sub-disciplines. They are not far behind PhD students in terms of using Mendeley. Whilst undergraduates are readers of scholarly articles in Mendeley, their scarcity compared to postgraduate and PhD students could be because undergraduates tend to use reference materials and textbooks (Jamali & Nicholas, 2006) rather than journal articles as they provide information in a more convenient way (Fescemyer, 2000). It could also be that they do not yet know about reference management software.

Many Clinical Medicine papers were read by people who are apparently not academics. This is important because some articles could be useful in clinical practice even if they are not cited in the literature (Jones, Donovan, & Hanney, 2012). Moreover, a substantial minority of the social science papers, probably Library and Information Science articles, had librarians as readers, which is consistent with results of Schloegl and Stock (2004). Additionally, librarians bookmarked some Clinical Medicine papers and this could be an indication of medical researchers engaging clinical librarians in scholarly activities like systematic searching and information dissemination. The importance of these kinds of scholarly activities by librarians has been mentioned before (Brettle & Long, 2001).

There were substantial and positive correlations between Mendeley readership and citation counts for all sub-disciplines studied: Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry. These findings corroborate previous studies (Bar-Ilan, 2012; Li et al., 2012) but stronger correlations were found in our dataset. One possible reason for the increased correlation is that the number of Mendeley users has grown over time, giving better raw readership data. As reported above, the correlations vary across different sub-disciplines.

The median Mendeley readership counts for all Social Science and some Engineering and Technology sub-disciplines are significantly higher than the median number of citations. These results suggest that papers in these research areas were bookmarked by many people that did not cite them, consistent with Social Science articles having many pure readers (Armbruster, 2008) and Engineering and Technology papers being used in applied contexts. Thus, it seems that Mendeley readership is able to provide evidence of using research articles in contexts other than for their science contribution, at least for Social Science and some applied sub-disciplines. Therefore, as citation-based indicators are less effective for social science research evaluation than for hard sciences research evaluation (Nederhof, 2006), Mendeley readership could compliment citations for the evaluation of social science articles. Moreover, it also could be used as a supplementary indicator to measure the impact of some technological or medical papers in applied contexts, as citation analysis is more useful for the assessment of theoretical research than applied research.

In response to the second research question, there are positive correlations between Mendeley readership and citations for all occupations except librarians for all of the sub-disciplines examined. However, the highest correlations are for users that are also authors, except for associate professors in some sub-disciplines. This suggests that counts of Mendeley readers with authorship roles probably reflect types of impact that are more similar to traditional citation impact in comparison to readership counts for non-author types of user. This goes some way towards validating Mendeley as an altmetric data source. Alternatively, it can be said that the highest correlations are for authors with the highest density of readerships in Mendeley. Nevertheless, the correlations for authors are not strong enough to claim that Mendeley readership counts and citation counts are interchangeable. It is likely that academics use research articles in activities other than citing, such as teaching. The lowest correlations were found for undergraduates and non-academic users. This suggests that students often benefit from articles that are not highly cited. Thus, Mendeley provides an opportunity to monitor impact on students, which probably reflects the educational value of research articles. This would only work for a small percentage of articles, however, since undergraduates are a small minority of Mendeley users and their data is typically hidden by Mendeley as a side effect of reporting only the three most common types of user for each article. A logical consequence of this is that low correlations for undergraduate users may partly be an artefact of readership counts for undergraduates only registering in the Mendeley API when they form a disproportionately high percentage of an article's readers. Similarly, non-academic readership counts have among the lowest correlations with citation counts, suggesting that their readership counts could also help to identify individual articles and types of article that are valuable outside academia.

The results of this study are consistent with the conclusions of Kurtz and Bollen (2010), which were based on case studies (e.g. Rowlands & Nicholas, 2007; Bollen & Van De Sompel, 2008) with usage datasets that were mainly local (institutional), publisher dependent and not publicly accessible, and therefore not practical for most researchers. Thus, Mendeley offers a practical source of global usage data for multiple disciplines for the first time.

One of the limitations of this study is that the sample is restricted to journal articles, while users in particular occupations or disciplines may benefit more from other document types. For example, engineers read a relatively high number of conference articles in comparison to books and journal papers (Niu & Hemminger, 2012).

Based on advanced Mendeley searches for WoS journal articles from 2008, it seems that Mendeley has records for 837,958 journal articles from 2008, although some of these records are likely to be duplicates and so the actual number is likely to be lower than this. Within these articles, 788,260 (94%) had at least one Mendeley reader. The main sample of the current study includes the 197,848 WoS articles from 2008 from Clinical Medicine, Engineering and Technology, Social Science, Physics and

Chemistry with at least one Mendeley reader, which is about 25% of the articles from 2008 with at least one Mendeley reader. The results may be different to some extent for the disciplines excluded, however.

Another limitation is that this study compares the individual readership counts of each article with the total citations received by it instead of the unique number of citing authors. The latter would involve author disambiguation, which was not feasible have given the large amount of citing papers.

As mentioned above, data on readers' occupations was only available for three most common reader categories for each article, which resulted in losing around 27% of the readership counts for some of the analyses. A consequence of this is that numerically small groups of readers (e.g., associate professors, professors, undergraduates, librarians) may have lower correlations due to underestimating the extent of their readership or only recording their readership values for articles for which they formed a disproportionately large share of the readers. Perhaps most importantly, Mendeley is most useful for those who will eventually cite an article and so its readership counts seem likely to under-represent users who will never need to cite an article, for example, disproportionately many practitioners. Hence, Mendeley readership statistics should not be taken as an unbiased reflection of an article's readers.

Finally, from the perspective of using Mendeley as a data source for altmetrics, the biggest limitation is that the users of Mendeley possibly form a small and biased minority of all readers of academic articles. In particular, assuming that Mendeley users tend to be younger than typical article readers, Mendeley readership data could not be used to estimate the proportions of different types of readers for articles. For example, although various types of professor form less than 10% of the Mendeley readers of articles and various types of student form 55-77% (depending on the area), it is possible that professors are the majority readers of articles but rarely join Mendeley. Nevertheless, it seems reasonable to compare the proportions of Mendeley readers with different occupations across different disciplines in Mendeley (e.g., Social Sciences vs. Clinical Medicine) to identify whether readership is particularly high for one occupational group, even though the level of uptake of Mendeley between different professions and academic position could also vary between disciplines.

CHAPTER 6: MOTIVATIONS FOR BOOKMARKING IN MENDELEY

6.1. Introduction

Researchers, science funders and evaluators may be interested to know who uses research outputs and to understand the contexts in which research is used. For instance, the ACUMEN Portfolio has proposed new indicators to assist research evaluators to avoid relying upon formal citations and expert judgements (see http://research-acumen.eu). One indicator of the use or value of an academic publication could be the extent of its readership within the scientific community, which gives evidence of interest in using it in research or other scholarly activities (e.g., teaching and discussions). Although getting data about the readership of academic publications can be difficult (Wouters & Costas, 2012), several studies have used download counts for electronic articles to indicate readership (Kurtz et al., 2005; Haque & Ginsparg, 2009). Nevertheless, it is not clear whether downloads can reflect something about intellectual impact. For example, the identities of downloaders of papers are unknown due to confidentiality and privacy issues, while the type of user can partly differentiate the contexts that articles are used for (Duin, King, & Van den Besselaar, 2012; Thelwall, 2012). In contrast, the social reference sharing site Mendeley provides data for scientific publications about the people that have saved information about each article to their Mendeley library (called 'bookmarking' in this chapter to distinguish it from reading, although this is not a strictly accurate term) *and* identifies their roles (e.g., professors, PhD students, masters students, and users outside academia). Mendeley bookmarking counts have been interpreted as an indicator of 'readership' (Bar-Ilan, 2012; Li, Thelwall, & Giustini, 2012; Thelwall, Haustein, Larivière, & Sugimoto, 2013; Zahedi, Costas, & Wouters, 2013) but the extent to which a bookmark represents a genuine reader is not known. In response, a large-scale online survey was conducted to directly investigate different aspects of using Mendeley. This seems to be the first large scale survey to investigate the reasons for using Mendeley to help interpret bookmarking counts.

The following research questions drive this investigation.

1. Why do users bookmark publications in Mendeley?
2. What are the main purposes for using Mendeley (e.g., reference manager, publicising publications, social networking)?
3. To what extent do Mendeley bookmarking counts reflect article readership?

6.2 Method and Procedure

In order to address the research questions, we needed to recruit an unbiased sample of Mendeley users. To do this, we required a comprehensive list of Mendeley users. Although Mendeley claims that it has more than 2.5 million users (Mendeley, 2013), a full list of its users is not available for researchers. However, Mendeley publishes a directory[13] which contained approximately 188,100 users from different disciplines at the time of data collection (October 2013). We used this directory as the largest available list of Mendeley users. Initially, we selected 5,000 random Mendeley users across all disciplines from the directory and tried to contact the chosen sample through sending direct messages via Mendeley messaging, with permission from

[13] http://www.mendeley.com/directory

Mendeley. This was not successful because, as an undocumented feature of Mendeley, only a limited number of messages can be sent each day.

Instead, we emailed Mendeley members using the contact information in their public home pages, if any. For this, a list of all users in the Mendeley directory was automatically extracted, including their research discipline and Mendeley profile URL. The distribution of all available Mendeley users in the directory across different disciplines is shown in Appendix 15. Next, using web searches in Webometric Analyst (lexiurl.wlv.ac.uk) with the query below, users who had personal web page URLs in their Mendeley profiles were identified automatically from the Bing (http://www.bing.com) API (Applications Programming Interface).

"Webpage:" "[two last keywords of the Mendeley Profile URL]" site:

http://www.mendeley.com/profiles

For example, the query `"Webpage:" "kayvan-kousha" site: http://www.mendeley.com/profiles/` captured the webpage www.koosha.tripod.com from the user's Mendeley profile. This process identified 19,959 users who had a URL "contact information" section in their Mendeley profile. Email addresses were manually collected from these webpages, when present, giving 6,122 for all disciplines. As shown in Appendix 15, the backgrounds of 8%, 23%, 10%, 26% and 26% of those in the Mendeley directory were arts and humanities, basic science, engineering, medicine and biology and social sciences, respectively. Similarly, 10%, 31%, 12%, 24% and 23% of the extracted emails belonged to users in arts and humanities, basic science, engineering, medicine and biology and social sciences, respectively. This means that the proportion of extracted email addresses for each subject area is only very approximately representative of the population of each discipline in the Mendeley directory. Detailed information about the number of Mendeley users with personal webpages in their ‘contact information’ and the extracted email addresses are reported in Appendix 15.

A questionnaire was designed to assess 1) the main reasons for using Mendeley 2) Mendeley users' motivations for bookmarking papers in their personal libraries, and 3) the proportion of bookmarked documents that Mendeley users had read or intend to read (see Appendix 16). The main themes of the questions for this survey originated from the need to find qualitative evidence for interpretations of Mendeley bookmarking counts to follow up chapters 3 and 4 of this thesis. The questionnaire was developed through a series of pilot tests with Mendeley users and altmetrics researchers. The survey received ethical approval from the University of Wolverhampton Research Institute for Information and Language Processing (see Appendix 17).

In the middle of January 2014, using Survey Monkey, email questionnaire invitations were sent to 5,927 Mendeley users (excluding invitations that bounced) across all disciplines (see Table 6.1). A reminder was sent to non-responding persons in late January. In order to improve the response rate, in the reminder we offered participants the chance to win one of ten $100 Amazon vouchers. As shown in Table 6.1, 14.6% (864) people responded to the survey, ranging from 13% in medicine and biology to 17% in the social sciences (Appendix 18). Altogether, 73% of the participants replied to the survey in the first call and 27% responded after the reminder.

Table 6.1. Email invitations sent to Mendeley users and response rates across disciplines.

Broad disciplines based on users' Mendeley profiles	Number (%) of users in the Mendeley directory	Number (%) of recruited users (excluding bounced emails)	Response rate (No.)
Arts and Humanities	14,380 (7.6%)	582 (9.8%)	13.1% (76)
Basic Science	43,727 (23.2%)	1,843 (31.1%)	13.7% (253)
Engineering	19,229 (10.2%)	7,11 (12%)	17.2% (122)
Medicine and Biology	48,881 (26.0%)	1,440 (24.3%)	12.9% (186)
Social Sciences	49,823 (26.5%)	1,351 (22.8%)	16.8% (227)
Total	188,100 (100%)	5,927 (100%)	14.6% (864)

6.3 Results

6.3.1 Occupation and Discipline of Respondents

Table 6.2 shows that over half of the survey respondents were PhD students (27%) or postdoctoral researchers (26%). Moreover, about 14%, 13% and 11% of the respondents were assistant professors, associate professors, and full professors respectively. About 6% of the Mendeley users in this survey were *other professionals*, which could mean that they were practitioners in a specific field, including engineers, surgeons and lawyers. The smallest categories were masters students (3%) and undergraduate students (1%).

Table 6.2. The occupations of the survey respondents (n=864).

User's Occupation	% of participants
PhD student	27%
Postdoctoral researcher	26%
Assistant professor/lecturer	14%
Associate professor/reader / senior lecturer	13%
Professor	11%
Other professionals	6%
Masters student	3%
Undergraduate student	1%

Table 6.3 shows that the proportions of survey participants for the social sciences and basic science were equal (27%). Around 26%, 16% and 6% of respondents to this survey defined their main discipline as engineering, medical sciences and arts and humanities respectively. Comparing Table 6.3 and Appendix 18, the respondents are approximately representative of the chosen sample subject areas in terms of numbers. The distribution of the participants is also in agreement with all Mendeley users at the level of broad disciplines to some extent (Mendeley, 2012).

Table 6.3. The broad subject areas of the survey respondents (n=864).

User's broad disciplines	% of participants
Basic science	27%
Social sciences	27%
Engineering	26%
Medical sciences	16%
Arts and humanities	6%

6.3.2 Motivations for Using Mendeley

About 78% of the respondents had a personal library in Mendeley, where they could potentially add other publications for different scholarly reasons (e.g., referencing and teaching) as evidence of readership. Most importantly, the majority of respondents (87%) reported that they used Mendeley as a reference manager, whereas only 30%, 25% and 15% used it as a database for searching academic publications, as a tool for publicising their publications, or as a social networking site.

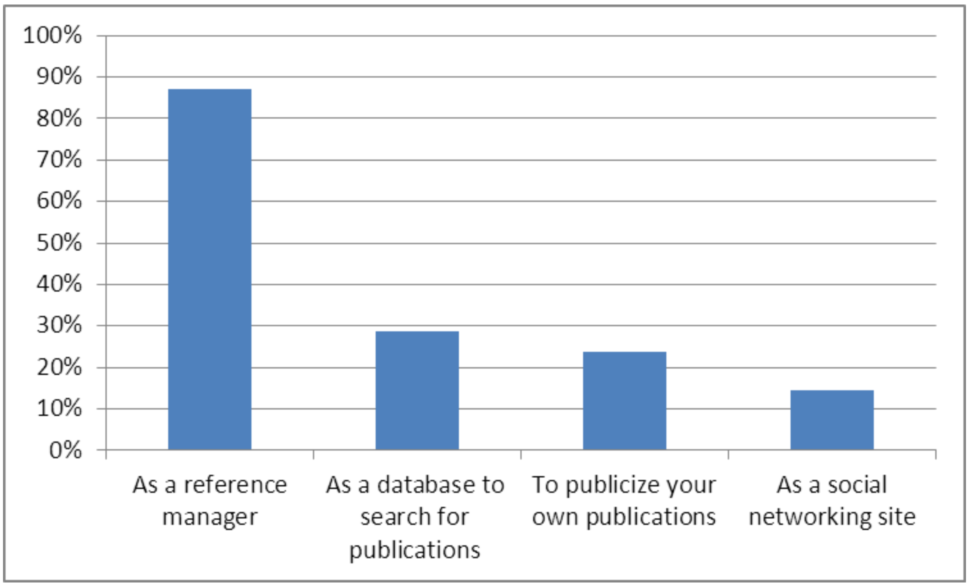

Figure 6.1. Purposes for using Mendeley, as reported by all survey respondents (n=864).

A chi-square test (p=0.691; see Table 6.4, Appendix 19) found no significant disciplinary differences in purposes for using Mendeley. Nevertheless, the significant chi-square value in Table 6.6 is probably due to the low proportion of Yes answers in Medical science (9%) compared to the average for the rest (17%). This suggests that medical science researchers are less likely to use Mendeley for publicity than are other researchers. Table 6.9 shows (Appendix 20) there were statistically significant differences in purposes for using Mendeley between users' occupations (p=0.025). For instance, academic staff (i.e., professors, associate professors, and assistant professors, researchers) used Mendeley to publicise their publications more than did the other professions and undergraduate students and masters students used Mendeley as a platform for searching academic publications more than did the other groups (see Table 6.9, Appendix 20).

6.3.3 Motivations for Bookmarking Papers in Personal Libraries

Disciplinary differences About 85% of the respondents across all disciplines bookmarked papers in Mendeley to cite them in their publications. There is strong evidence (p=0.001, see Table 6.14, Appendix 21) that *overall* motivations for bookmarking papers differ between disciplines.

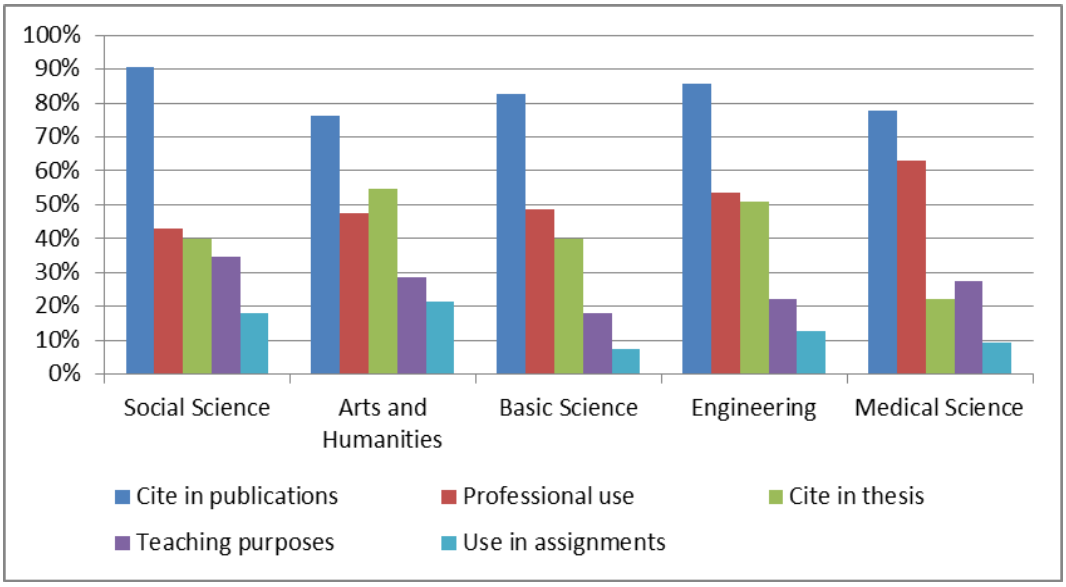

Figure 6.2. Mendeley users' motivations for bookmaking papers in their personal library by discipline (n=679).

Although bookmarking with the aim of *future citation* was the most common reason across all disciplines, the percentage varied from 91% for social science users to 76%

for arts and humanities users, and these differences are statistically significant (p=0.023, see Table 6.15).

The second most common purpose for adding scholarly publications to Mendeley was for *professional use* in all subject areas (50%). About 63% and 54% of Mendeley users in medical science and engineering bookmarked academic records for professional use while the proportion decreased to 49%, 48% and 43% for users with basic science, art and humanities and social science backgrounds, respectively. Table 6.19 reports that there were statically significant differences between the disciplines (p=0.021) in terms of bookmarking papers in Mendeley for professional use.

Approximately 40% of the participants in all disciplines added records to their Mendeley libraries to *cite in their thesis or dissertation* and this percentage was 55% for arts and humanities and 51% for engineering. Approximately 25% of Mendeley users bookmarked publications in their teaching activities in all disciplines. Only 13% of users bookmarked documents in Mendeley to use them in assignments and these were presumably all students.

Occupation differences Figure 5.3 indicates that the most common reason for bookmarking scientific publications for professors (83%), associate professors (88%), assistant professors (94%), and PhD students (88%) was to cite in their publications. Unsurprisingly, there were statistically significant differences in motivations for bookmarking documents in Mendeley between different user occupations (p=0.000, Table 6.20). More specifically, the aims of bookmarking papers for future citation, teaching and educational activities and professional use also differed between occupations (Table 6.29, 5.30 and 5.31). Figure 5.3 also shows that around 85% of both PhD students and masters students added documents to their Mendeley personal libraries to cite in their theses. Similarly, 38% of masters students and 20% of PhD students bookmarked documents in their personal collections for course assignments. Additionally, the aim of bookmaking scientific documents in Mendeley for teaching activities was more common for professors (45%), associate professors (36%) and assistant professors (33%) than for the other groups.

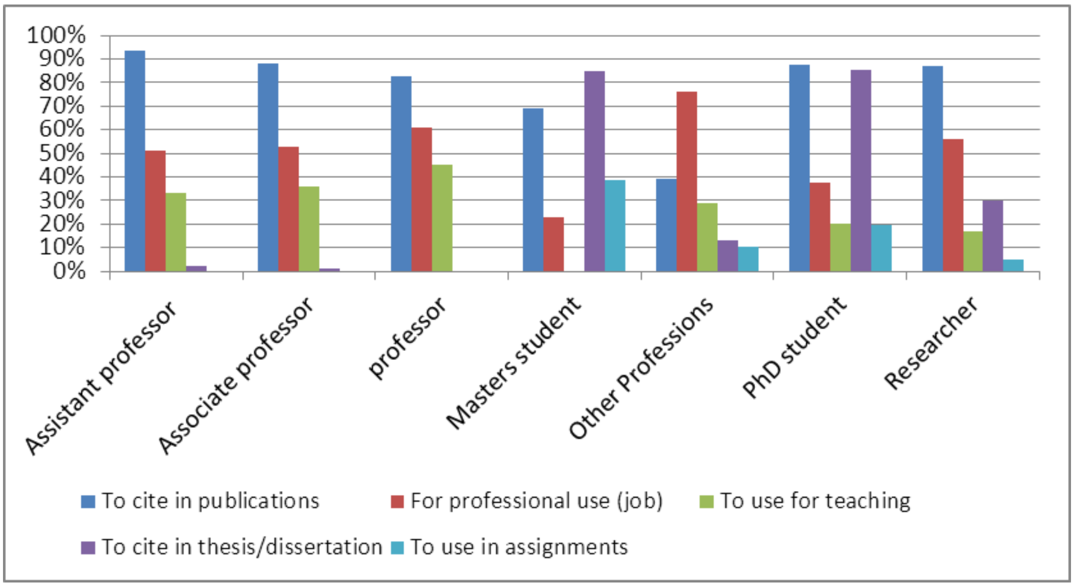

Figure 6.3. Mendeley users' motivations for bookmaking papers in their personal libraries based on occupation (n=679).

Figure 5.3 shows that the principal purpose for bookmarking academic publications for Mendeley users with *other professional* backgrounds was to use these documents in their professional activities (76%) (e.g., medical activities), although 39% bookmarked documents to cite in their publications and so were active in research to some extent.

6.3.4 Reading Bookmarked Publications

A total of 679 out of 864 respondents had a personal library in Mendeley. Moreover, 27% of users with a personal library in Mendeley had read all of their bookmarked records, 55% had read at least half and 18% had read less than half of the bookmarked items. Almost none (0.4%) of the users did not read the bookmarked records. In total, 82% of the Mendeley users had read at least half of the bookmarked publications in their personal libraries. A chi square test (p=0.282, see Appendix 23) found no significant disciplinary differences in the proportions of the items from personal libraries that the survey respondents had read.

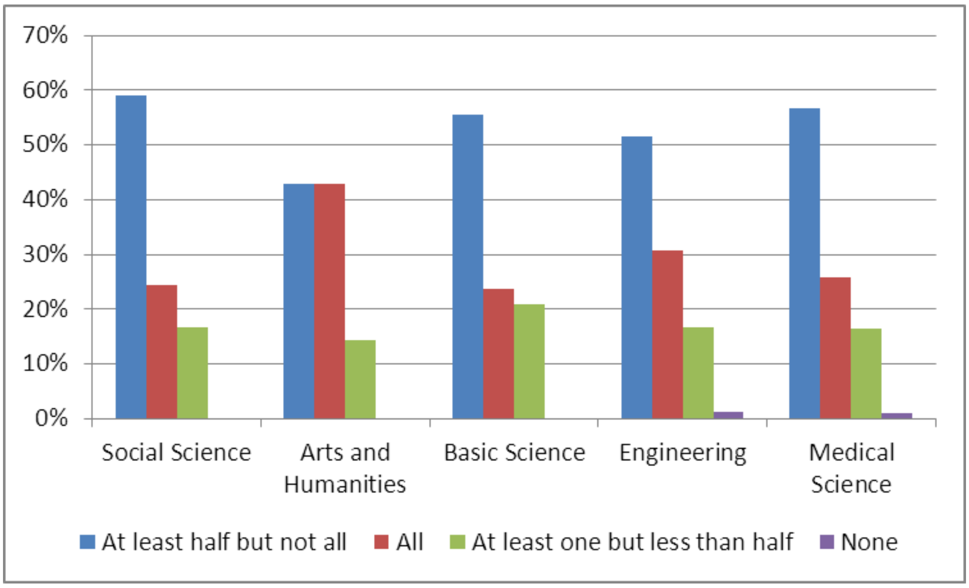

Figure 6.4. The proportion of the items from Mendeley personal libraries that survey respondents had read by discipline (n=679).

6.4 Discussion

Users both within and outside of academia participated in the survey. Most were PhD students and postdoctoral researchers. This is in agreement to some extent with the quantitative results from chapter 4 and previous findings (Schloegl, Gorraiz, Gumpendorfer, Jack, & Kraker, 2013; Zahedi, Costas and Wouters, 2013) that most readers of publications in Mendeley were PhD students or postdoctoral researchers. However, the number of masters students was low in this survey in comparison with the results of chapter 4, which reported that masters students were among most common readers of Mendeley papers. A possible reason for this is the lower visibility of contact information for masters students in Mendeley in comparison with other categories (e.g., if fewer listed a personal webpage). In other words, survey respondents were representative of readers of articles in terms of occupations to some extent, except for masters students. Another likely reason that most survey respondents were PhD students and postdoctoral researchers is that Mendeley is a new platform and motivates younger researchers to use it, while senior scholars are unlikely to use many social web platforms (Mas-Bleda, Thelwall, Kousha, & Aguillo, in press).

The survey respondents were representative of the initial survey sample in terms of academic disciplines and to some extent of all Mendeley users (Mendeley, 2012). However, the method of categorising academic disciplines in Mendeley and in this survey may not be identical and so this is a tentative conclusion.

In response to the first research question, the most common reason for bookmarking publications in Mendeley was to cite them in future publications. This finding corroborates the results of Tenopir, King, Spencer, and Wu (2009) which indicated that the principal motivation for reading academic publications is for use in future research, although their sample excluded professional users outside universities. Mendeley users who were authors of scholarly publications (e.g., professors, assistant professors) were most likely to bookmark papers for future citation. However, it is possible that people who do not author papers can also cite papers in non-journal publications (e.g., dissertations). Therefore, Mendeley bookmarking counts can partly represent future citations and this is consistent with the medium correlations between citations and Mendeley bookmarking counts in chapters 3 and 4 and previous studies (Bar-Ilan, 2012; Li, Thelwall, & Giustini, 2012; Thelwall, Haustein, Larivière, & Sugimoto, 2013; Zahedi, Costas, & Wouters, 2013).

Interestingly, around half of the Mendeley users bookmarked publications for professional use, and this amount was higher for those with backgrounds in applied disciplines, such as medical science and engineering. Simiarly, the main motivation for *other professionals* (e.g., engineers, surgeons and lawyers) for bookmarking papers in Mendeley seemed to be related to their professional activities because they work outside academia. This is evidence that Mendeley could be used to track the use of academic publications in practical contexts.

The survey found that the majority of masters students and PhD students with a personal library in Mendeley bookmarked academic publications for citing them in their thesis. This is a good method of capturing academic publications that have been used in theses, as few seem to be covered by citation databases such as WoS or Scopus. This means that Mendeley readership counts may be able to capture citation-like activities in broader contexts than those covered by conventional citation indexes. Moreover, a substantial minority of masters students (38%) and PhD students (20%) bookmarked records for completing their assignments, reflecting the educational value of these publications. A high proportion of professors (45%), associate professors (36%) and assistant professors (36%) added records to their Mendeley library for use in their teaching activities. This was particularly true for users with social science backgrounds (see Table 6.25 and 5.34) which is in agreement with the findings of Tenopir et al. (2009). In summary, the current study suggests that Mendeley bookmarking counts reflect several scholarly activities, including future citation in publications and theses, for use in practical contexts, and for application in teaching and educational activities.

In answer to the second research question, the study showed that the main reason for using Mendeley was for managing references (87%). This is not surprising as the main aim of Mendeley is providing facilities to manage references. Respondents of this survey also used Mendeley for academic literature searching (30%). Publicising their own publications was the third most common reason (25%) for using Mendeley but it was rarely used as a social network site. Mendeley was perhaps not popular as a social network site because some of its social features are not free. For instance, free plan Mendeley users can only create one private group with up to three members. Nevertheless, although the main aim of using Mendeley was to organise and archive references, it was also used for other purposes.

Most survey respondents (79%) had a Mendeley personal library and only a small proportion of them had not bookmarked any publications. Users without a personal library may join Mendeley for other purposes, such as increasing the visibility of their own publications through their Mendeley profiles, or they may have created an account in Mendeley without actively using it.

In response to the third research question, the survey indicates that most Mendeley users read most of the bookmarked publications in their personal libraries. This provides direct evidence that Mendeley bookmark counts can reflect readership, although not that Mendeley bookmarking counts are proportional to the number of readers of a publication, due to the sampling representativeness issues discussed above.

6.5. Limitations

The findings are subject to a number of limitations. First, the relatively low response rate to the survey may influence the results, presumably by under-representing less enthusiastic Mendeley users and particularly busy scientists. Second, the Mendeley directory was the source of the sampled Mendeley users but it covers only 188,100 out of the 2.5 million Mendeley users and there is no information about the criteria for listing users in this directory. Thus, the representativeness of the initial sample is unknown, although it broadly matches some known properties of the Mendeley user base. Third, the sample recruited in this survey is limited to Mendeley users mentioning their personal websites in their Mendeley profiles, and this is likely to bias the findings towards Mendeley users with a greater web presence, disadvantaging masters and undergraduate students. Additionally, the Bing API used to capture the personal webpages in Mendeley user profiles has unknown coverage of this site. Hence, the findings should be only cautiously generalised to all Mendeley users. Finally, although this study revealed that the majority of Mendeley users had read or believed that they would eventually read the bookmarked records in their personal libraries, this perception may not be true, and does not reflect how carefully articles are read.

6.6 Conclusions

This study confirms that different types of users within academia, from undergraduate students to full professors, use Mendeley to bookmark academic publications. Moreover, professionals outside academia also bookmark scholarly publications in Mendeley but, unsurprisingly, there were relatively few of them. This study also confirms that Mendeley is mainly used to manage references (87%).

Nevertheless, Mendeley is also used for purposes beyond reference managing, such as searching for scholarly documents (30%), publicising publications (25%), and networking with peers (15%).

The survey findings suggest that Mendeley bookmarking counts may be a useful measure of readership because most records that were bookmarked had been read or were planned to be read. The possibility of connecting bookmarking records to attributes of the readers (i.e., profession, discipline) means that Mendeley can help to reveal information about readers of academic papers, in contrast to typical download data (Moed, 2005). Nevertheless, Mendeley bookmarking cannot necessarily reflect the full spectrum of types of reader of academic articles because there may be some kinds of reader do not use Mendeley.

Although the most common motivation for bookmarking documents in Mendeley was for citing in future research, the findings provide evidence of the use of academic publications for professional use by both scholars within academia and professionals outside academia, especially for some applied fields such as medical science and engineering. Additionally, documents are sometimes also bookmarked in Mendeley for other purposes, such as for use in teaching and education. Hence, Mendeley bookmarking counts can capture a variety of contexts of use for academic publications. The reasons for the previously discovered significant moderate correlations between Mendeley readership and citations (Li, Thelwall, & Giustini, 2011; Bar-Ilan, 2012; Zahedi, Costas, & Wouters, 2013) may be due to the main motivations for bookmarking documents in Mendeley, and the correlations are perhaps not strong because of the variety of purposes for bookmarking papers in Mendeley. For example, some of the highly bookmarked papers may be useful in education rather than research.

In summary, this study confirms that Mendeley bookmarking counts reflect an aspect of the readership of scholarly publications and can partly reflect some types of research impact beyond traditional citation impact.

CHAPTER 7: CONCLUSIONS

7.1 Introduction

This doctoral thesis originated from the need to assess new metrics for research evaluation using two social web sites: Mendeley and F1000. This is part of the broader altmetrics aim to develop new approaches for measuring invisible research impact within and outside academia. This thesis investigates the validity of the new metrics through comparisons with citation counts as recognised indicators of research impact. This study also examines the types of research impact that can be identified through the new metrics and the extent to which the new indicators can be used as complementary metrics to help overcome drawbacks of citations for research assessment. This thesis also explores disciplinary differences for the new metrics. This project contributes to current scientific knowledge by meeting the objectives of the original investigations as follows.

7.2 Assessing Non-standard Article Impact Using F1000 Labels

The F1000 study suggests that some types of research can be identified by reviewers as being, on average, more valuable than others in ways that are not recognised through citations. In particular, F1000 scores could, in theory, be used in research evaluation exercises when the importance of practical findings needs to be recognised due to better highlighting of papers with "Changes to Clinical Practice" through FFa scores. Furthermore, since the majority of the articles studied were reviewed in their publication year, F1000 can also be useful for early impact evaluations to overcome the slowness of citation. A problem that would need to be addressed, however, is how such evaluations should deal with articles that have not been reviewed in F1000."

7.3 Examining correlations between Mendeley readership counts and citations

Chapters 4 and 5 found statistically significant medium positive correlations between Mendeley readership counts and citations across all disciplines, although the correlations varied between disciplines (tables 4.2 and 5.3). In all disciplines the correlations between Mendeley readership counts and citations are not strong enough to claim that these two metrics measure the same aspects of research impact. These findings are in agreement with some recent studies (Bar-Ilan, 2012; Li, Thelwall, & Giustini, 2012; Thelwall, Haustein, Larivière, & Sugimoto, 2013).

The median Mendeley readership counts were higher than the median citations for the articles covered by Mendeley in some of the disciplines (see tables 3.2 and 4.3) particularly for social sciences, humanities and engineering. Additionally, a significant minority of papers without citations have Mendeley readers and the proportions of these papers are highest in the Social Sciences (15%) and Engineering and Technology (11%) (see appendices 24-28). The five most read papers with zero citations in Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry are listed in Appendix 29 as evidence. Clearly, some papers are extensively bookmarked by Mendeley users despite not being cited in WoS. This gives evidence that Mendeley readership data could help to identify the impact of articles that appear from WoS data to have had no impact, especially in social sciences, humanities and engineering. However, it would be useful to explore papers with high Mendeley readership and low citations in depth in future studies. If Mendeley readership data is to be used for important evaluations, however, then steps would need to be taken to ensure that the results cannot be manipulated by those with a vested interest in a particular outcome. This finding now suggests that Mendeley may be universally useful for research impact estimation throughout all areas of scholarship. This adds to the previously-known fact that Mendeley has the advantage of covering broader types of users while citation data comes only from authors. Mendeley data may also appear earlier than citations because of the lack of publication delays.

7.4 Investigating Information Flow among Disciplines Based on Mendeley Readers

Although citation analysis is often used to map knowledge transfer between disciplines, its limitations (see section 2.1.2) mean that it not ideal for revealing knowledge flows between scientific fields. In this thesis, Mendeley readership data are is also used to map knowledge transfer between different disciplines. Comparing Mendeley readership data with citation analysis (see section 3.3.2) confirms that Mendeley data give similar results and hence is a reasonable alternative for discovering information flows between scientific disciplines. Mendeley readership data will be particularly useful when recent information flows are needed or a wider perspective than just publishing authors is needed.

7.5 Identifying Readers of Scholarly Publications Based on Professions

This thesis found that the majority of readers for all disciplines were PhD students, postgraduates and postdocs but other types of academics were also represented. In terms of Mendeley readers outside of higher education institutions, these appear to be a small minority, with Clinical Medicine having 7.2% from “Other Professions” (for papers with 100% readership counts). Thus, Mendeley readership is able to capture a dimension of the impact of scientific documents on various activities performed within the academic community such as “plain reading” (i.e., reading without subsequently citing, writing theses, doing assignments or drafting research proposals) but also provides a little evidence of their applied use by people outside academia, such as medical doctors and surgeons.

7.6 Investigating the Effect of Academic Status on the Correlations Between Citations and Mendeley Bookmarking Counts

Mendeley readership counts could perhaps supplement citation counts in the social sciences and in some engineering research areas in which citation counts are lower than Mendeley readership counts. The variation in correlations between Mendeley readership counts and citations received for different types of reader suggest that the meaning of Mendeley readership counts depends upon the readers’ occupations. In some cases Mendeley readership may reflect traditional citation impact but in other cases it may reflect educational uses or impact on applied contexts. Therefore, Mendeley readership data is different from both citations and raw usage data. However, Mendeley is only one of many reference manager tools and other reference managers (e.g., Endnote, RefWorks, Zotero) also have many users but their data are not publically available. Thus, Mendeley seems to be the only choice to reveal aspects of the readership of research articles but it can only reveal part of that readership, even considering only public reference manager users. It could be particularly useful in disciplines for which citation-based indicators are least reliable, such as the social sciences, arts and humanities, and perhaps for applied research.

7.7 Establishing Motivations for Bookmarking Scholarly Publications in Mendeley

This thesis confirmed that different types of users within academia, from undergraduate students to full professors, use Mendeley to record academic publications. Moreover, professionals outside academia also bookmark scholarly publications in Mendeley but, unsurprisingly, they did this considerably less often than did users inside academia. This thesis also found that Mendeley users typically join this platform in order to manage academic publications, which is not surprising because the stated aim of Mendeley is to provide services for managing references. Mendeley users also used it for other purposes, such as searching for scholarly documents, publicizing their own publications, and networking with peers.

The results of the survey suggest that it is reasonable to use Mendeley readership counts as an indicator of readership because users tend to record articles that they have read or intend to read. Nevertheless, it is not reasonable to claim that Mendeley readership can represent all readership because it seems likely that only a tiny minority of article readers record them in Mendeley.

This thesis suggests that there are many motivations for recording documents in Mendeley. Unsurprisingly, the primary motivation is for citing in future research. Additionally, several other reasons for recording articles in Mendeley were identified, such as use in teaching and educational activities. Hence, Mendeley readership counts can capture different ways using academic publications. This explains the low and medium correlations between citations and Mendeley readership counts of previous studies (Li, Thelwall, & Giustini, 2011; Bar-Ilan, 2012; Zahedi, Costas, & Wouters, 2013) and the findings of chapters 3 and 4 of this thesis. In other words, the reason for significant correlations between Mendeley readership and citations may be that the main motivations for bookmarking documents in Mendeley relate to referencing whereas the weak correlations could be due to the variety of purposes for bookmarking papers in Mendeley.

In summary, this thesis confirms that Mendeley readership counts are useful for capturing aspects of the readership of scholarly publications and are useful altmetrics for identifying different aspects of the research impact of articles.

7.8 Future Research

Altmetrics need to be assessed from both quantitative and qualitative perspectives. From the quantitative point of view, more studies are needed to explore different document types in different social platforms because document type importance varies by discipline. For example, the book is important in social science and humanities disciplines and conference papers are important in engineering. Additionally, due to the differences in the levels of activities of scholars across different disciplines, future studies need to find a way to suggest normalized altmetrics for academic fields similar to the normalized citation indicators. Future altmetrics investigations also need to go beyond quantitative approaches and discover the meaning of numbers behind altmetrics counts. Although this thesis partly tackles this issue with an online survey of Mendeley users, further research is need to find out who are the users of different social platforms and their motivations for using these tools in their scholarly activities. Due to the unregulated nature of the social web, altmetrics are not as sustainable as citation data and future studies need to find solutions to minimize the risk posed by this. To implement altmetrics in practice researchers and developers in altmetrics need to collaborate to find practical, meaningful and sustainable metrics for different types of research evaluation.

REFERENCES

Abbott, A., Cyranoski, D., Jones, N., Maher, B., Schiermeier, Q., & Van Noorden, R. (2010). Do metrics matter? *Nature*, 465(7300), 860–862.

Advisory Board for the Research Councils (ABRC). (1990). Peer Review. A Report to the Advisory Board for the Research Councils from the Working Group on Peer Review. London.

Aksnes, D. W., & Taxt, R. E. (2004). Peer reviews and bibliometric indicators: a comparative study at a Norwegian university. *Research Evaluation*, 13(1), 33–41.

Allen, L., Jones, C., Dolby, K., Lynn, D., & Walport, M. (2009). Looking for landmarks: the role of expert review and bibliometric analysis in evaluating scientific publication outputs. *PloS one*, 4(6), e5910.

Almind, T. C., & Ingwersen, P. (1997). Informetric analyses on the world wide web: methodological approaches to "webometrics". *Journal of Documentation*, 53(4), 404–426.

Armbruster, C. (2008). Access, Usage and Citation Metrics: What Function for Digital Libraries and Repositories in Research Evaluation? *Social Science Research Network* Working Paper Series, 128(Pt 6), 1407–17.

Andersen, J. (2013). Conceptualising research quality in medicine for evaluative bibliometrics. forskningsbasen. *Unpublished doctoral dissertation*, University of Copenhagen. Copenhagen, Denmark.

Bar-Ilan, J. (2012). JASIST 2001–2010. *Bulletin of Association for Information Science and Technology*, 24–28.

Barjak, F., & Thelwall, M. (2008). A statistical analysis of the web presences of European life sciences research teams. *Journal of the American Society for Information Science and Technology*, 59(4), 628–643.

Barnett, G. A., & Fink, E. L. (2008). Impact of the internet and scholar age distribution on academic citation age. *Journal of the American Society for Information Science and Technology*, *59*(4), 526–534.

Becher, T., & Trowler, P. (2001). Academic tribes and territories (2ed). Milton Keynes, UK: Open University Press.

Bennett, N. L., Casebeer, L. L., Kristofco, R. E., & Strasser, S. M. (2004). Physicians' Internet information-seeking behaviors. *The Journal of continuing education in the health professions*, 24(1), 31–8.

Björneborn, L. (2004). Small-world link structures across an academic Web space: A library and information science approach. *Doctoral dissertation*, Royal School of Library and Information Science, Copenhagen, Aalborg, Denmark.

Björneborn, L., & Ingwersen, P. (2004). Toward a basic framework for webometrics. *Journal of the American Society for Information Science and Technology*, 55(14), 1216–1227.

Blecic, D. D. (1999). Measurements of journal use: an analysis of the correlations between three methods. *Bulletin of the Medical Library Association*, 87(1), 20–5.

Boell, S. K., & Wilson, C. S. (2010). Journal Impact Factors for evaluating scientific performance: use of h-like indicators. *Scientometrics*, 82(3), 613–626.

Bollen, J, Rodriquez, M., & Sompel, H. Van de. (2006). Journal status. *Scientometrics*, 69, (3) 669–687.

Bollen, J., & Sompel, H. V. (2008). Usage impact factor: The effects of sample characteristics on usage based impact metrics. *Journal of the American Society for Information Science and Technology*, 59(1), 136–149.

Bollen, J., Van de Sompel, H., Hagberg, A., & Chute, R. (2009). A principal component analysis of 39 scientific impact measures. (T. Mailund, Ed.)*PloS one*, 4(6), e6022.

Bollen, J., Van de Sompel, H., Smith, & Luce. (2005). Toward alternative metrics of journal impact: A comparison of download and citation data. *Information Processing & Management*, 41(6), 1419–1440.

Bordons, M., Morillo, F., & Gómez, and I. (2005). analysis of cross-disciplinary research through bibliometric tools. In H. F. Moed (Ed.), *Handbook of Quantitative Science and Technology Research* (pp. 437–456).

Bornmann, L. (2010). Mimicry in science? *Scientometrics*, 86(1), 173–177.

Bornmann, L. (2012). Measuring the societal impact of research: research is less and less assessed on scientific impact alone--we should aim to quantify the increasingly important contributions of science to society. *EMBO Reports*, 13(8), 673–6.

Bornmann, L, & Daniel, H. (2005). Selection of research fellowship recipients by committee peer review. Reliability, fairness and predictive validity of Board of Trustees' decisions. *Scientometrics*, 63(2), 297–320.

Bornmann, L. (2014). Is there currently a scientific revolution in scientometrics? *Journal of the Association for Information Science and Technology,* 65: 647–648.

Bornmann, L., & Daniel, H.-D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, 64(1), 45–80.

Bornmann, L., & Leydesdorff, L. (2012). The validation of (advanced) bibliometric indicators through peer assessments: A comparative study using data from InCites and F1000. *Arxiv preprint arXiv*: 1211.1154.

Bornmann, L. and Marx, W. (2014), The wisdom of citing scientists. *Journal of the Association for Information Science and Technology*, 65: 1288–1292

Borrego, Á., & Fry, J. (2012). Measuring researchers' use of scholarly information through social bookmarking data: A case study of BibSonomy. *Journal of Information Science*, 38(3), 297-308.

Braun, T. (2012). Special Discussion Issue on Journal Impact Factors [Special issue]. *Scientometrics*, 92(2), 207–208.

Brettle, A J., & Long, A F. (2001). Comparison of bibliographic databases for information on the rehabilitation of people with severe mental illness. *Bulletin of the Medical Library Association*, 89(4), 353–62.

Brody, T., Harnad, S., & Carr, L. (2006). Earlier Web Usage Statistics as Predictors of Later Citation Impact. *Journal of the American Society for Information Science and Technology*, 57(8), 1060–1072.

Buschman, M., & Michalek, A. (2013). Are alternative metrics still alternative? *Bulletin of Association for Information Science and Technology*, 35–39.

Butler, L. (2007). Assessing university research: a plea for a balanced approach. *Science and Public Policy*, 34(8), 565–574.

Camacho-Miñano, M.-M., & Núñez-Nickel, Manuel. (2009). The multilayered nature of reference selection. *Journal of the American Society for Information Science and Technology*, 60(4), 754–777.

Catalano, A. (2013). *Patterns of graduate students' information seeking behavior: a meta-synthesis of the literature*. *Journal of Documentation* (Vol. 69, pp. 243–274).

Chu, H., Krichel, T., & Blvd, N. (2007). Downloads vs . Citations: Relationships , Contributing Factors and Beyond 1 Palmer School of Library & Information Science , Long Island University ,. In H.F. Moed (Eds.) (Ed.), *Proceedings of the 11th International Society for Scientometrics and Informetrics Conference* (pp. 207–215). Madrid: CINDOC.

Costas, R., Zahedi, Z., & Wouters, P. (2014). Do altmetrics correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Arxiv preprint arXiv*: 1401.4321.

Cronin, B. (1984). The citation process. The role and significance of citations in scientific communication. London: Taylor Graham.

Cronin, B. (2005). The Hand of Science: Academic Writing and Its Rewards (p. 214). Lanham, MD: Scarecrow Press.

Cronin, B. (2013a). Metrics à la mode. *Journal of the American Society for Information Science and Technology*, 64(6), 1091–1091.

Cronin, B. (2013b). The evolving indicator space (iSpace). *Journal of the American Society for Information Science and Technology*, 64(8), 1523–1525.

Cronin, B. (2014). Scholars and Scripts, Spoors and Scores. In B. Cronin & C. R. Sugimoto (Eds.), Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact (p. 432). MIT Press.

Cronin, B., & Meho, L. (2007). The shifting balance of intellectual trade in information studies. *Journal of the American Society for Information Science and Technology*, *59*(4), 551–564.

Cronin, B., & Pearson, S. (1990). The export of ideas from information science. *Journal of Information Science*, *16*(6), 381–391.

Cronin, B., Snyder, H., & Atkins, H. (1997). Comparative citation rankings of authors in monographic and journal literature: a study of sociology. *Journal of Documentation*, 53(3), 263–273.

Cronin, B., Snyder, H. W., Rosenbaum, H., Martinson, A., & Callahan, E. (1998). Invoked on the Web. *Journal of the American Society for Infor- mation Science*, 49(14), 1319–1328.

Darmoni, S., & Roussel, F. (2002). Reading factor: a new bibliometric criterion for managing digital libraries. *Journal of the Medical Library*, 90(3), 323–327.

De Jong, S. P. L., Van Arensbergen, P., Daemen, F., Van der Meulen, B., & Van den Besselaar, P. (2011). Evaluation of research in context: an approach and two cases. *Research Evaluation*, 20(1), 61–72.

Dellavalle, R. P., Schilling, L. M., Rodriguez, M. A., Van de Sompel, H., & Bollen, J. (2007). Refining dermatology journal impact factors using PageRank. *Journal of the American Academy of Dermatology*, 57(1), 116–9.

Drooge, L., Besselaar, P., Elsen, G., & Haas, M. (2010). Evaluating the societal relevance of academic research: A guide. The Hague: Rathenau Instituut.

Duin, D., King, D., & Van den Besselaar, P. (2012). Identifying audiences of e-infrastructures - tools for measuring impact. *PloS one*, 7(12), e50943.

Duy, J., & Vaughan, L. (2006). Can electronic journal usage data replace citation data as a measure of journal use? An empirical examination. *The Journal of Academic Librarianship*, 32(5), 512–517.

Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics*, 69(1), 131–152.

Egghe, L. (1988). Mathematical relations between impact factors and average number of citations. *Information Processing & Management*, 24(5), 567–576.

Egghe, Leo, & Rousseau, R. (1990). Introduction to Informetrics: quantitative methods in library, documentation and information science. Amsterdam: Elsevier Science Publishers.

Ellis, D., & Haugan., M. (1997). Modelling the information seeking patterns of engineers and research scientists in an industrial environment. *Journal of documentation*, *4*(53), 384–403.

Eysenbach, G. (2011). Correction: Can Tweets Predict Citations? Metrics of Social Impact Based on Twitter and Correlation with Traditional Metrics of Scientific Impact. *Journal of Medical Internet Research*, 13(4), e123.

F1000. (2012). F1000 Faculty. Retrieved from http://f1000.com/prime/thefaculty

F1000. (2014). Article ratings and rankings. Retrieved fromhttp://f1000.com/prime/about/whatis/factors

Falagas, M. E., Kouranos, V. D., Arencibia-Jorge, R., & Karageorgopoulos, D. E. (2008). Comparison of SCImago journal rank indicator with journal impact factor. FASEB journal: official publication of the Federation of American Societies for Experimental Biology, 22(8), 2623–8.

Fenner, M. (2013). What Can Article-Level Metrics Do for You? *PLoS Biology*, 11(10).

Fenner, M. (2014). Altmetrics and Other Novel Measures for Scientific Impact. In S. Bartling & S. Friesike (Eds.), Opening Science (pp. 179–189). Springer International Publishing.

Fescemyer, K. (2000). Information-seeking behavior of undergraduate geography students. *Research Strategies*, *17*(4), 307–317.

Fienberg, S. E., & Martin, M. E. (1985). Sharing research data. Washington:The National Academies Press.

Franceschet, M., & Costantini, A. (2011). The first Italian research assessment exercise: a bibliometric perspective. *Journal of Informetrics*, 5(2), 275-291.

Freund, L., Toms, E. G., & Waterhouse, J. (2005). Modeling the Information Behaviour of Software Engineers Using a Work - Task Framework.

Proceedings of the American Society for Information Science and Technology, *42*(1).

Garfield, E. (2011). Full Text downloads and citations: Some reflections. Keynote lecture at the Seminar “Scientific Measurement and Mapping”. Santa Fe, New Mexico.

Ge, X. (2010). Information-seeking behavior in the digital age: A multidisciplinary study of academic researchers. *College & Research Libraries*, 435–455.

Gingras, Y., Larivière, V., Macaluso, B., & Robitaille, J.-P. (2008). The effects of aging on researchers’ publication and citation patterns. *PloS one*, 3(12), e4048.

Goldstone, R. L., & Leydesdorff, L. (2006). The import and export of cognitive science. *Cognitive science*, *30*(6), 983–93.

Hahn, K., & Faulkner, L. (2002). Evaluative Usage-based Metrics for the Selection of E-journals. *College & Research Libraries*, 215–227.

Hanney, S., Frame, I., Grant, J., Buxton, M., Young, T., & Lewison, G. (2005). Using categorisationsof citations when assessing the outcomes from health research. *Scientometrics*, 65(3), 357–379.

Haque, A., & Ginsparg, P. (2009). Positional effects on citation and readership in arXiv. *Journal of the American Society for Information Science and Technology*, 60(11), 2203–2218.

Harnad, S. (2008). Open access scientometrics and the UK Research Assessment Exercise. *Scientometrics*, 79(1), 147–156.

Haustein, S. (2013). Altmetrics: Present and Future; a Panel discussion at the ASIS&T 2013 Annual Meeting.Montreal, Canada.

Haustein, S., Golov, E., & Luckanus, K. (2010). Journal evaluation and science 2.0: Using social bookmarks to analyze reader perception. *Book of Abstracts of the 11th International Conference on Science and Technology Indicators* (pp. 117–119).

Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2013). Coverage and adoption of altmetrics sources in the bibliometric community. *Arxiv preprint arXiv*: 1304.7300

Haustein, S., Peters, I., Sugimoto, C. R., Thelwall, M., & Larivière, V. (In press). Tweeting biomedicine: An analysis of tweets and citations in the biomedical literature. *Journal of the Association for Information Science and Technology.*

Haustein, S., & Siebenlist, T. (2011). Applying social bookmarking data to evaluate journal usage. *Journal of Informetrics*, 5(3), 446–457.

HEFCE. (2011). Research Excellence Framework (Vol. 2011). Higher Education Funding Council for England. http://www.ref.ac.uk/media/ref/content/pub/decisionsonassessingresearchimpact/01_11.pdf

Hemminger, B. M., Lu, D., Vaughan, K. T. L., & Adams, S. J. (2007). Information seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology*, 58(14), 2205–2225.

Hertzum, M., & Pejtersen, A. M. (2000). The information-seeking practices of engineers: searching for documents as well as for people. *Information Processing & Management*, *36*(5), 761–778.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United states of America, 102(46), 16569.

Holbrook, J. B. (2010). The use of societal impacts considerations in grant proposal peer review: a comparison of five models. *Technology & Innovation*, 12(3), 213–224.

Hood, W., & Wilson, C. (2001). The literature of bibliometrics , scientometrics , and informetrics. *Scientometrics*, 52(2), 291–314.

Huggett, S. (2012). F1000 Journal Rankings: an alternative way to evaluate the scientific impact of scholarly communications. *Research Trends*, (26), 7–11.

Ingwersen, P. (1998). The calculation of web impact factors. *Journal of Documentation*, 54(2), 236–243.

Iorns, E. (2012). Reproducibility: An important altmetric. altmetrics12: ACM Web Science Conference 2012 Workshop. Evanston, IL.

Jamali, H. R., & Nicholas, D. (2006). Communication and information-seeking behavior of PhD students in physicists and astronomy. *Proceedings of the American Society for Information Science and Technology*, *43*(1), 1–18.

Jansen, B. J. (2006). Search log analysis: What it is, what's been done, how to do it. *Library & Information Science Research*, 28(3), 407–432.

Jayasinghe, U. W., Marsh, H. W., & Bond, N. (2003). A multilevel cross-classified modelling approach to peer review of grant proposals: the effects of assessor and researcher attributes on assessor ratings. *Journal of the Royal Statistical Society*: Series A (Statistics in Society), 166(3), 279–300.

Jiang, J., He, D., & Ni, C. (2011). Social Reference: Aggregating Online Usage of Scientific Literature in CiteULike for Clustering Academic Resources. *JCDL'11 June 13–17 2011* (Vol. 11, pp. 401–402). Ottawa, Ontario, Canada: ACM.

Jones, T. H., Donovan, C., & Hanney, S. (2012). Tracing the wider impacts of biomedical research: a literature search to develop a novel citation categorisation technique. *Scientometrics,* 1–10.

King, D., & Tenopir, C. (2004). An evidence-based assessment of the author-pays model. Nature web focus: Access to the literature.

Kamalski, J., Plume, A., & Mayur, A. (2014). A Publishing Perspective on Bibliometrics. In E. Garfield, B. Cronin, & H. B. Atkins (Eds.), *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact* (p. 432). MIT Press.

Konkiel, S. (2013). Altmetrics: A 21st-Century Solution to Determining Research Quality. *Online searcher,* (July/August).

Korobili, S., Malliari, A., & Zapounidou, S. (2011). Factors that Influence Information-Seeking Behavior: The Case of Greek Graduate Students. *The Journal of Academic Librarianship*, *37*(2), 155–165.

Kostoff, R. (2007). The difference between highly and poorly cited medical articles in the journal Lancet. *Scientometrics*, 72(3), 513–520.

Kostoff, R. N. (1998). The use and misuse of citation analysis in research evaluation. *Scientometrics*, 43(1), 27–43.

Kostoff, R. N. (1997). The principles and practices of peer review. *Science and Engineering Ethics*, 3(1), 19–34.

Kousha, K., & Thelwall, M. (2007a). Google Scholar citations and Google Web/URL citations: A multi-discipline exploratory analysis. *Journal of the American Society for Information Science and Technology*, 58(7), 1055–1065.

Kousha, K., & Thelwall, M. (2007b). How is science cited on the web? A classification of Google unique web citations. *Journal of the American Society for Information Science and Technology,* 58(11), 1631–1644.

Kousha, K., & Thelwall, M. (2008). Assessing the impact of disciplinary research on teaching: An automatic analysis of online syllabuses. *Journal of the American Society for Information Science and Technology*, 59(13), 2060–2069.

Kousha, K., & Thelwall, M. (2009). Google book search: Citation analysis for social science and the humanities. *Journal of the American Society for Information Science and Technology,* 60(18),1537-1549

Kousha, K., & Thelwall, M. (2014). Web Impact Metrics for Research Assessment. In C. R. Cronin, Blaise Sugimoto (Ed.), Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact (p. 342). MIT Press.

Kousha, K., Thelwall, M., & Rezaie, S. (2010). Can the impact of scholarly images be assessed online? An exploratory study using image identification technology.

Journal of the American Society for Information Science and Technology, 61(9), 1734–1744.

Kousha, K., Thelwall, M., & Rezaie, S. (2011). Assessing the citation impact of books: The role of Google Books, Google Scholar, and Scopus. *Journal of the American Society for Information Science and Technology,* 62(11), 2147–2164.

Kraker, P., Körner, C., Jack, K., & Granitzer, M. (2012). Harnessing User Library Statistics for Research Evaluation and Knowledge Domain Visualization. *Proceedings of the 21st international conference companion on World Wide Web WWW 12 Companion* (p. 1017). ACM.

Kurtz, M. J., & Bollen, J. (2010). Usage bibliometrics. *Annual Review of Information Science and Technology*, 44(1-64).

Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C., Demleitner, M., Murray, S. S., Martimbeau, N., et al. (2005). The bibliometric properties of article readership information. *Journal of the American Society for Information Science and Technology*, 56(2), 111–128.

Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C. S., Murray, S. S., & Watson, J. M. (2000). The NASA Astrophysics Data System: Overview. *Astronomy and Astrophysics Supplement Series*, 143(1), 41–59.

Kwasitsu, L. (2003). Information-seeking behavior of design , process , and manufacturing engineers. *Library & Information Science Research*, *25*(4), 459–476.

Larivière, V., & Gingras, Y. (2010). On the relationship between interdisciplinarity and scientific impact. *Journal of the American Society for Information Science and Technology*, *61*(1), 126–131.

Larivière, V., Sugimoto, C. R., & Bergeron, P. (2013). In their own image? a comparison of doctoral students' and faculty members' referencing behavior. *Journal of the American Society for Information Science and Technology*, 64(5), 1045–1054.

Leckie, G., Pettigrew, K., & Sylvain, C. (1996). Modeling the information seeking of professionals: a general model derived from research on engineers, health care professionals, and lawyers. *The Library Quarterly*.

Levitt, J., & Thelwall, M. (2011). Variations between subjects in the extent to which the social sciences have become more interdisciplinary. *Journal of the American Society for Information Science and Technology*, *62*, 1118–1129.

Lewison, G. (2002). From biomedical research to health improvement. *Scientometrics*. 54(2), 179-192.

Lewison, G., & Sullivan, R. (2008). The impact of cancer research: how publications influence UK cancer clinical guidelines. *British Journal of Cancer*, 98(12), 1944–1950.

Leydesdorff, L. (2009). How are new citation-based journal indicators adding to the bibliometric toolbox? *Journal of the American Society for Information Science and Technology*, 60(7), 1327–1336.

Li, & Thelwall, M. (2012). F1000 , Mendeley and Traditional Bibliometric Indicators. *17th International Conference on Science and Technology Indicators* (Vol. 3, pp. 1–11).

Li, X., Thelwall, M., & Giustini, D. (2012). Validating online reference managers for scholarly impact measurement. *Scientometrics*, 91(2), 461–471.

Li, X., Thelwall, M., Wilkinson, D., & Musgrove, P. (2005). National and international university departmental Web site interlinking. *Scientometrics*, 64(2), 151–185.

Lin, J., & Fenner, M. (2013). Altmetrics in Evolution: Defining and Redefining the Ontology of Article-Level Metrics. *Information Standards Quarterly*, 25(2), 20.

Liu, J., & Adie, E. (2013). Five challenges in altmetrics: A toolmaker's perspective. *Bulletin of the american society for information science and technology*, 31–34.

Lockett, A., & McWilliams, A. (2005). The Balance of Trade Between Disciplines: Do We Effectively Manage Knowledge? *Journal of Management Inquiry*, *14*(2), 139–150.

McAlister, F. A., Graham, I., Karr, G. W., & Laupacis, A. (1999). Evidence-based medicine and the practicing clinician. *Journal of General Internal Medicine*, *14*(4), 236–242.

MacRoberts, M. H., & MacRoberts, B. R. (1989). Problems of citation analysis: A critical review. *Journal of the American Society for Information Science*, 40(5), 342–349.

MacRoberts, M. H., & MacRoberts, B. R. (1996). Problems of citation analysis. *Scientometrics*, 36(3), 435–444.

MacRoberts, M. H., & MacRoberts, B. R. (2010). Problems of citation analysis: A study of uncited and seldom-cited influences. *Journal of the American Society for Information Science and Technology*, 61(1), 1–12.

Mahdi, S., D'Este, P., & Neely, A. D. (2008). Citation counts: are they good predictors of RAE scores?: a bibliometric analysis of RAE 2001. AIM Research.

Maier, G. (2006). Impact factors and peer judgment: The case of regional science journals. *Scientometrics*, 69(3), 651–667.

Martin, B. R. (1996). The use of multiple indicators in the assessment of basic research. *Scientometrics*, 36(3), 343–362.

Mas Bleda, A., Thelwall, M., Kousha, K. & Aguillo, I. (2013). Do Highly Cited Researchers Successfully use the Social Web. *Scientometrics*..

Mendeley (2012). Mendeley Has 2.5 Million Users. Retrieved February 20, 2014 from http://blog.mendeley.com/start-up-life/mendeley-has-2-5-million-users/

Mendeley (2012). Mendeley Global Research Report. Retrieved February 20, 2014 from http://www.mendeley.com/global-research-report/#.UPxyUqPi58E

Merkx, F., Weijden, I. van der, Oostveen, A.-M., Besselaar, P. van den, & Spaapen, J. (2007). Evaluation of research in context: A quick scan of an emerging field (second edition). Rathenau Institute, Department of Science System Assessment (p. 34). The Hague:

Mervis, J. (2011). Beyond the data. *Science (New York, N.Y.)*, 334(6053), 169–71.

Meyer, M. (2002). Tracing knowledge flows in innovation systems. *Scientometrics*, 54(2), 193–212.

Moed, H.F. (2005). Citation analysis in research evaluation (Vol. 9). Kluwer Academic Pub.

Moed, H F. (2005). Statistical relationships between downloads and citations at the level of individual documents within a single journal. *Journal of the American Society for Information Science and Technology*, 56(10), 1088–1097.

Moed, H. F. (2007). The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer review. *Science and Public Policy*, 34(8), 575–583.

Moed, H F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, 4(3), 265–277.

Moed, H F, Glänzel, W., & Schmoch, U. (2004). Handbook of quantitative science and technology research: the use of publication and patent statistics in studies of S & T systems. Springer.

Moed, H.F., Leeuwen, T. N. Van, & Reedijk, J. (1998). A new classification system to describe the ageing of scientific journals and their impact factors. *Journal of Documentation*, 54(4), 387–419.

Moed, H. F., Colledge, L., Reedijk, J., Moya-Anegon, F., Guerrero-Bote, V., Plume, A., & Amin, M. (2012). Citation-based metrics are appropriate tools in journal assessment provided that they are accurate and used in an informed way. *Scientometrics*, 92(2), 367–376.

Moed, H. F., Linmans, J., & Nederhof, A. (2009). Options for a comprehensive database of research outputs in Social Sciences & Humanities, (April).

Morillo, F., Bordons, M., & Gómez, I. (2003). Interdisciplinarity in science: A tentative typology of disciplines and research areas. *Journal of the American Society for Information Science and Technology*, *54*(13), 1237–1249.

Narin, F. (1976). Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity. Washington, D. C: Computer Horizons.

Nederhof. (2006). Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review. *Scientometrics*, 66(1), 81–100.

Nederhof, A. J., & Van Raan, A. F. J. (1993). A bibliometric analysis of six economics research groups: A comparison with peer review. *Research Policy*, 22(4), 353–368.

Neeley, J. D. (1981). The management and social science literatures: An interdisciplinary cross-citation analysis. *Journal of the American Society for Information Science*, *32*(3), 217–223.

Neylon, C., & Wu, S. (2009). Article-Level Metrics and the Evolution of Scientific Impact. *PLoS Biology*, 7(11), 6.

Nicholas, D., Huntington, P., Dobrowolski, T., Rowlands, I., Jamali M., H. R., & Polydoratou, P. (2005). Revisiting "obsolescence" and journal article "decay" through usage data: an analysis of digital journal use by year of publication. *Information Processing & Management*, 41(6), 1441–1461.

Niu, X., & Hemminger, B. M. (2012). A study of factors that affect the information-seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology*, 63(2), 336–353.

Norris, M., & Oppenheim, C. (2003). Citation counts and the Research Assessment Exercise V: Archaeology and the 2001 RAE. *Journal of Documentation*, 59(6), 709–730.

NSF. (2013). Proposal and award policies and procedures guide; Changes to the Grant Proposal Guide (GPG) Effective January 14, 2013. Virginia.

Oppenheim, C. (1995). The correlation between citation counts and the 1992 Research Assessment Exercise Ratings for British library and information science university departments. *Journal of Documentation*, 51(1), 18–27.

Oppenheim, C., & Summers, M. A. C. (2008). Citation counts and the Research Assessment Exercise, part VI: Unit of assessment 67 (music). *Information Research*, 13(2), 3.

Opthof, T. (1997). Sense and nonsense about the impact factor. *Cardiovascular Research*, 33(1), 1–7.

Opthof, T., & Leydesdorff, L. (2011). A comment to the paper by Waltman et al., Scientometrics, 87, 467–481, 2011. *Scientometrics*, 88(3), 1011–1016.

Pair, C. (1980). Switching between academic disciplines in universities in the Netherlands. *Scientometrics*, *2*(3), 177–191.

Pancheshnikov, Y. (2007). A Comparison of Literature Citations in Faculty Publications and Student Theses as Indicators of Collection Use and a Background for Collection Management at a University Library. *The Journal of Academic Librarianship*, *33*(6), 674–683.

Pierce, S. J. (1999). Boundary crossing in research literatures as a means of interdisciplinary information transfer. *Journal of the American Society for Information Science*, *50*(3), 271–279.

Piwowar, H. (2013). Altmetrics: Value all research products. *Nature*, 493(7431), 159–159.

Piwowar, H. A., Day, R. S., & Fridsma, D. B. (2007). Sharing detailed research data is associated with increased citation rate. *PloS one*, 2(3).

Price, D. D. S., & Gürsey, S. (1975). Studies in Scientometrics I Transience and Continuance in Scientific Authorship. Ciência da Informação, 4(1).

Priem, J, Groth, P., & Taraborelli, D. (2012). The Altmetrics Collection. (C. A. Ouzounis, Ed.) *PLoS ONE*, 7(11).

Priem, J, & Hemminger, B. M. H. (2010). Scientometrics 2.0: New metrics of scholarly impact on the social Web. *First Monday*, 15(7).

Priem, J., Parra, C., Piwowar, H., Groth, P., & Waagmeester, A. (2012). Uncovering impacts: a case study in using altmetrics tools. Workshop on the Semantic Publishing SePublica 2012 at the 9th Extended Semantic Web Conference.

Priem, J. (2013). Altmetrics and Revolutions:New Products, New Metrics, and the New Era of Web-native Scholarship. SLA Annual Conference. San Diego, CA, USA.

Priem, J. (2014). Altmetrics. In C. R. Cronin, Blaise Sugimoto (Ed.), Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact (p. 342). MIT Press.

Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. *Arxiv preprint arXiv*:1203.4745.

Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2011). altmetrics: a manifesto.

Qiu, J., Chen, J., & Wang, Z. (2004). An analysis of backlink counts and Web Impact Factorsfor Chinese university websites. *Scientometrics*, 60(3), 463–473.

Reale, E., Barbara, A., & Costantini, A. (2007). Peer review for the evaluation of academic research: lessons from the Italian experience. *Research Evaluation*, 16(3), 216–228.

Rinia, E. . (2007). *Measurement and evaluation of interdisciplinary research and knowledge transfer*. Universiteit Leiden, The Netherlands.

Rinia, E. J., Van Leeuwen, T. N., Bruins, E. E. W., Van Vuren, H. G., & Van Raan, A. F. J. (2001). Citation delay in interdisciplinary knowledge exchange. *Scientometrics*, 51(1), 293–309.

Rinia, E. J., Van Leeuwen, T., Bruins, E., Van Vuren, H., & Van Raan, A. (2002). Measuring knowledge transfer between fields of science. *Scientometrics*, 54(3), 347–362.

Rousseau, R. (2000). Exponential decline in the use distribution of medical journals. *Journal of Documentation*, 56(4), 454–5.

Rousseau, R. (1997). Sitations: an exploratory study. *Cybermetrics*, 1(1), 1–7.

Rousseau, R, & Ye, F. (2013). A multi-metric approach for research evaluations. *Chinese Science Bulletin*, 1–7.

Rowlands, I, & Nicholas, D. (2007). The missing link: journal usage metrics. *Aslib Proceedings*, 59(3), 222–228.

Rowlands, I, & Nicholas, D. (2005). Scholarly communication in the digital environment: The 2005 survey of journal author behaviour and attitudes. *Aslib Proceedings*, 57(6), 481–497.

Schilling, L. M., Steiner, J. F., Lundahl, K., & Anderson, R. J. (2005). Residents' patient-specific clinical questions: opportunities for evidence-based learning. *Academic medicine : journal of the Association of American Medical Colleges*, *80*(1), 51–6.

Schloegl, C., & Gorraiz, J. (2010). Comparison of citation and usage indicators: the case of oncology journals. *Scientometrics*, 82(3), 567–580.

Schloegl, C., & Gorraiz, J. (2011). Global usage versus global citation metrics: The case of pharmacology journals. *Journal of the American Society for Information Science and Technology*, 62(2005), 161–170.

Schloegl, C., & Stock, W. G. (2004). Impact and relevance of LIS journals: A scientometric analysis of international and German-language LIS journals - Citation analysis versus reader survey. *Journal of the American Society for Information Science and Technology*, 55(13), 1155–1168.

Schloegl, C., Gorraiz, J., Gumpendorfer, C., Jack, K., & Kraker, P. (2013). Download vs. Citation vs. Readership Data: The Case of an Information Systems Journal. 14th International Society of Scientometrics and Informetrics Conference. (pp. 626 – 634). Vienna: AustrianInstitute of Technology.

Seng, L. B., & Willett, P. (1995). The citedness of publications by United Kingdom library schools. *Journal of Information Science*, 21(1), 68–71.

Shema, H., Bar-Ilan, J., & Thelwall, M. (2013). Do blog citations correlate with a higher number of future citations? Research blogs as a potential source for alternative metrics. *Journal of the American Society for Information Science and Technology*, 52900(2003), 1–20.

Smith, A. (1998). A Tale of Two Web Spaces: Comparing Sites Using Web Impact Factors. *Journal of Documentation*, 55(5), 577–92.

Smith, A. T., & Eysenck, M. (2002). The correlation between RAE ratings and citation counts in Psychology. London.

Smith, R. (2006). Peer review: a flawed process at the heart of science and journals. *Journal of the Royal Society of Medicine*, 99(4), 178–82.

Sridhar, M. S. (1990). A Study of Correlation of Use, Citation and Publishing of Journal Papers by Indian Space Technologists. *Collection Management*, 12(1-2), 147–152.

Stern, R. E. (1990). Uncitedness in the biomedical literature. *Journal of the American society for information science*, 41(3), 193–196.

Stevens, G. (1990). An Alliance Confirmed Planning Literature and the Social Sciences. *Journal of the American Planning Association*, *56*(3), 341–349.

Sud, P., & Thelwall, M. (2014). Evaluating altmetrics. *Scientometrics*, 98(2), 1131-1143.

Sutherland, W. J., Goulson, D., Potts, S. G., & Dicks, L. V. (2011). Quantifying the impact and relevance of scientific research. *PloS One*, 6(11), e27537.

Tague-Sutcliffe, J. (1992). An introduction to informetrics. *Information Processing & Management*, 28(1), 1–3.

Tang, R. (2005). Evolution of the interdisciplinary characteristics of information and library science. *Proceedings of the American Society for Information Science and Technology*, *41*(1), 54–63.

Tenopir, C. (2003). “Use and users of electronic library resources: an overview and analysis of recent research studies”. Report for the Council on Library and Information Resources.

Tenopir, C., & King, D. W. (2000). Towards Electronic Journals: Realities for Scientists, Librarians, and Publishers. Washington, D. C.: Special Libraries Association.

Tenopir, C., King, D. W., Clarke, M. T., Na, K., & Zhou, X. (2007). Journal reading patterns and preferences of pediatricians. *Journal of the Medical Library Association : JMLA*, *95*(1), 56–63.

Tenopir, C., King, D. W., Spencer, J., & Wu, L. (2009). Variations in article seeking and reading patterns of academics: What makes a difference? *Library & Information Science Research*, *31*(3), 139–148.

Thelwall, M. (2008). Bibliometrics to webometrics. *Journal of Information Science*, 34(4), 605–621.

Thelwall, M. (2011). A comparison of link and URL citation counting. *Aslib Proceedings*.63(4), 419-425.

Thelwall, M. (2012). Journal impact evaluation: a webometric perspective. *Scientometrics*, 92(2), 429–441.

Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do Altmetrics Work? Twitter and Ten Other Social Web Services. (Lutz Bornmann, Ed.) *PLoS ONE*, 8(5), e64841.

Thelwall, M. and Kousha, K. (2014), Academia.edu: Social network or Academic Network?. *Journal of the Association for Information Science and Technology*, 65: 721–731

Thelwall, M., & Kousha, K. (2008). Online presentations as a source of scientific impact? An analysis of PowerPoint files citing academic journals. *Journal of the American Society for Information Science and Technology*, 59(5), 805–815.

Thelwall, M., & Maflahi, N. (2014). Are scholarly articles disproportionately read in their own country? An analysis of Mendeley readers. *Journal of the Association for Information Science and Technology*. doi: 10.1002/asi.23252

Thelwall, M., Vaughan, L., & Björneborn, L. (2006). Webometrics. *Annual Review of Information Science and Technology*, 39(1), 81–135.

Van Raan, A F J. (2006). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics*, 67(3), 491–502.

Van Raan, A. F. J. (2000). The Pandora's Box of Citation Analysis: Measuring Scientific Excellence - the Last Evil? In E. Garfield, B. Cronin, & H. B. Atkins (Eds.), The Web of Knowledge: A Festschrift in Honor of Eugene Garfield (pp. 301–320). Medford, New Jersey: ASIS Monograph Series: Information Today.

Van Raan, A.F.J. (1988). Handbook of Quantitative Studies of Science and Technology (p. 784). Elsevier Science Ltd.

Vaughan, L., & Hysen, K. (2002). Relationship between links to journal Web sites and impact factors. *Aslib Proceedings*. 54(6), 356-361.

Vaughan, L., & Shaw, D. (2003). Bibliographic and web citations: what is the difference? *Journal of the American Society for Information Science and Technology*. 54(14),1313-1322.

Vaughan, L., & Shaw, D. (2005). Web citation data for impact assessment: A comparison of four science disciplines. *Journal of the American Society for Information Science and Technology*, 56(10), 1075–1087.

Vaughan, L., & Thelwall, M. (2003). Scholarly use of the Web: What are the key inducers of links to journal Web sites? *Journal of the American Society for Information Science and Technology*, 54(1), 29–38.

Veletsianos, G., & Kimmons, R. (2012). Assumptions and challenges of open scholarship. *The International Review Of Research In Open And Distance Learning*, 13(4), 166–189.

Urata, H. (1990). Information flows among academic disciplines in Japan. *Scientometrics*, *18*(3-4), 309–319.

Waltman, L, Van Eck, N. J., Van Leeuwen, T. N., Visser, M. S., & Van Raan, A. F. J. (2011). On the correlation between bibliometric indicators and peer review: reply to Opthof and Leydesdorff. *Scientometrics*, 1–6.

Waltman, L. and Costas, R. (2014), F1000 Recommendations as a Potential New Data Source for Research Evaluation: A Comparison With Citations. *Journal of the Association for Information Science and Technology*, 65: 433–445.

Waltman, Ludo, Van Eck, N. J., Van Leeuwen, T. N., Visser, M. S., & Van Raan, A. F. J. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, 5(1), 37–47.

Wan, J., Hua, P., Rousseau, R., & Sun, X. (2010). The journal download immediacy index (DII): experiences using a Chinese full-text database. *Scientometrics*, 82(3), 555–566.

Wardle, D. A. (2010). Do'Faculty of 1000'(F1000) ratings of ecological publications serve as reasonable predictors of their future impact? *Ideas in Ecology and Evolution*, 3(0).

Whitmire, E. (2002). Disciplinary differences and undergraduates' information-seeking behavior. *Journal of the American Society for Information Science and Technology*, *53*(8), 631–638.

Wouters, P. (1997). Citation cycles and peer review cycles. *Scientometrics*, 38(1), 39–55.

Wouters, P., & Costas, R. (2012). Users , narcissism and control – tracking the impact of scholarly publications in the 21 st century. (M. Van Berchum & K. Russell, Eds.)Image Rochester NY, 50 pages.

Wulff, J. L., & Nixon, N. D. (2004). Quality markers and use of electronic journals in an academic health sciences library. *Journal of the Medical Library Association: JMLA*, 92(3), 315–22.

Zahedi, Z., Costas, R., & Wouters, P. (2013). How well developed are Altmetrics? Cross-disciplinary analysis of the presence of “alternative metrics” in scientific publications. *14th International Society of Scientometrics and Informetrics Conference (ISSI)* (pp. 876–884). Vienna: Austrian Institute of Technology.

Zahedi, Z., Costas, R., & Wouters, P. (2013). What is the impact of the publications read by the different Mendeley users? Could they help to identify alternative types of impact? *PLoS ALM Workshop*. San Francisco, US. https://www.academia.edu/6670997/What_is_the_impact_of_the_publications_read_by_the_different_Mendeley_users_Could_they_help_to_identify_alternative_types_of_impact and https://openaccess.leidenuniv.nl/handle/1887/23579

APPENDICES

Appendix 1: Faculty members of F1000 in different disciplines of medical sciences.

Topic	Associated and faculty members	Articles associated with the topic	Articles reviewed per faculty member
Anesthesiology and Pain Management	447	4803	11
Cardiovascular Disorders	156	3787	24
Critical Care andEmergency Medicine	141	3712	26
Dermatology	221	2591	12
Diabetes and Endocrinology	229	3744	16
Gastroenterology and Hepatology	296	3869	13
Hematology	238	2633	11
Infectious Diseases	253	4962	20
Nephrology	149	2002	13
Neurological Disorders	413	6183	15
Oncology	172	5371	31
Ophthalmology	175	536	3
Otolaryngology	149	1479	10
Psychiatry	180	3544	20
Public Health and Epidemiology	100	5314	53
Research Methodology	41	690	17
Respiratory Disorders	231	3520	15
Rheumatology and Clinical Immunology	313	3259	10
Urology	152	2397	16
Women's Health	124	4231	34
Total	4180	68627	

Appendix 2: Complete list of all merged Mendeley categories for readers of Social Sciences articles.

Merged category	Original Mendeley category	Total readership
Psychology	Psychology	35.04%
Social Sciences	Sports and Recreation	0.52%
	Social Sciences	10.22%
	Law	0.15%
Education	Education	7.60%
Business and Economics	Economics	7.06%
	Business Administration	13.96%
Management Science and Operations Research	Management Science Operations Research	4.35%
Computer and Information Science	Computer and Information Science	6.65%
Medicine	Medicine	4.33%
Biological Sciences	Biological Sciences	4.18%
Philosophy	Philosophy	0.56%
Linguistics	Linguistics	1.26%
Arts and Humanities	Arts and Literature	0.24%
	Humanities	0.47%
Others	Physics	0.16%
	Mathematics	0.21%
	Materials Science	0.02%
	Environmental Sciences	1.02%
	Engineering	1.16%
	Electrical and Electronic Engineering	0.15%
	Earth Sciences	0.16%
	Design	0.28%
	Chemistry	0.10%
	Astronomy, Astrophysics and Space Science	0.01%
Total		234231

Appendix 3: Complete list of all merged categories for readers of humanities articles.

Merged category	Original Mendeley category	Total readership
Philosophy	Philosophy	7.08%
Humanities	Humanities	12.62%
Linguistics	Linguistics	29.30%
Arts and Literature	Design	0.17%
	Arts and Literature	4.87%
Social Sciences	Sports and Recreation	0.05%
	Social Sciences	15.82%
	Management Science and Operations Research	0.37%
	Law	0.38%
Psychology	Psychology	10.17%
Education	Education	5.95%
Business and Economics	Economics	0.86%
	Business Administration	0.60%
Medicine	Medicine	1.26%
Biological Sciences	Biological Sciences	1.70%
Computer and Information Science	Computer and Information Science	7.24%
Others	Physics	0.24%
	Mathematics	0.10%
	Materials Science	0.00%
	Environmental Sciences	0.48%
	Engineering	0.35%
	Electrical and Electronic Engineering	0.12%
	Earth Sciences	0.13%
	Chemistry	0.02%
	Astronomy, Astrophysics and Space Science	0.01%
Total		7286

Appendix 4: Complete list of all merged WoS subject categories for disciplines citing Social Sciences articles.

Merged Category	**ISI Subject Category**
Psychology	Psychology
Social sciences	Criminology Penology
	Family Studies
	Biomedical Social Sciences
	Government Law
	Sociology
	Communication
	Social Work
	Women Studies
	Social Issues
	Anthropology
	Information Science Library Science
	Mathematical Methods In Social Sciences
	International Relations
	Ethnic Studies
	Cultural Studies
	Urban Studies
	Area Studies
Education educational research	Education Educational Research
Business and Economics	Business and Economics
Operations research management science	Operations Research Management Science
	Public Administration
Medicine and health	Psychiatry
	Public Environmental Occupational Health
	Rehabilitation
	Pediatrics
	Sport Sciences
	Pharmacology Pharmacy
	General Internal Medicine
	Health Care Sciences Services
	Radiology Nuclear Medicine Medical Imaging
	Nutrition Dietetics
	Substance Abuse
	Nursing
	Geriatrics Gerontology
	Ophthalmology
	Audiology Speech Language Pathology
	Oncology

Merged Category	**ISI Subject Category**
	Endocrinology Metabolism
	Obstetrics Gynecology
	Research Experimental Medicine
	Otorhinolaryngology
	Urology Nephrology
	Anesthesiology
	Surgery
	Cardiovascular System Cardiology
	Respiratory System
	Immunology
	Infectious Diseases
	Rheumatology
	Orthopedics
	Integrative Complementary Medicine
	Veterinary Sciences
	Hematology
	Medical ethics
	Dentistry Oral Surgery Medicine
	Gastroenterology Hepatology
	Toxicology
	Dermatology
	Demography
	Allergy
	Legal Medicine
	Tropical Medicine
	Emergency Medicine
	Medical Laboratory Technology
Biology and Life Sciences	Physiology
	Life Sciences Biomedicine Other Topics
	Genetics Heredity
	Environmental Sciences Ecology
	Zoology
	Biochemistry Molecular Biology
	Evolutionary Biology
	Cell Biology
	Biophysics
	Developmental Biology
	Reproductive Biology
	Anatomy Morphology
	Biotechnology Applied Microbiology
	Pathology
	Virology
	Plant sciences
	Biodiversity Conservation

Merged Category	ISI Subject Category
	Marine Freshwater Biology
	Microbiology
	Parasitology
Computer Science	Computer Science
Engineering	Engineering
Linguistics	Linguistics
Art and Humanities	Philosophy
	Music
	Arts Humanities Other Topics
	Religion
	History Philosophy Of Science
	Literature
	History
	Film Radio Television
	Archaeology
	Dance
	Art
Others	Science Technology Other Topics
	Mathematics
	Transportation
	Food science technology
	Medical Informatics
	Acoustics
	Agriculture
	Chemistry
	Mathematical Computational Biology
	Physics
	Robotics
	Materials Science
	Optics
	Telecommunications
	Imaging Science Photographic Technology
	Instruments Instrumentation
	Automation Control Systems
	Transplantation
	Architecture
	Construction Building Technology
	Electrochemistry
	Meteorology Atmospheric Sciences
	Energy Fuels
	Nuclear Science Technology
	Polymer Science
	Physical Geography
	Geography

Appendix 5: Complete list of all merged WoS subject categories for disciplines citing humanities articles.

Merged Category	**ISI Subject Category**
Philosophy	Philosophy
	History Philosophy of Science
History	History
	Archaeology
Linguistics	Linguistics
Literature	Literature
Religion	Religion
Arts Humanities Other Topics	Arts Humanities Other Topics
	Theater
	Film Radio Television
	Classics
	Art
	Music
Social Sciences	Cultural studies
	Social Sciences Other Subjects
	Communication
	Anthropology
	Area Studies
	Women Studies
	Information Science Library Science
	International Relations
	Asian Studies
	Urban Studies
	Ethnic Studies
	Public Administration
	Biomedical Social Sciences
	Social Work
Psychology	Psychology
Education Educational Research	Education Educational Research
Business and Economics	Business and Economics
Medicine and Health	Public Environmental Occupational Health
	Psychiatry
	Cardiovascular System Cardiology
	Neurosciences Neurology
Biology and Life Sciences	Immunology
	Biochemistry Molecular Biology
Computer Science	Computer Science
Others	Materials Science
	Science Technology Other Topics
	Chemistry
	Physics
	Agriculture
	Food Science Technology
	Geography

Appendix 6: Coverage of WoS articles from Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry in Mendeley (detailed version).

Sub-discipline	Articles indexed by WoS in 2008	Unique WoS articles covered by Mendeley	Duplicated records in Mendeley	Articles with readership statistics in Mendeley	Articles without readership statistics	Total readership including duplicates	Lost readership counts after removing duplicates
Neurology and Neurosurgery	31,616	77.8%	0.0%	73.0%	4.7%	288730	0.3%
Pharmacology	23,276	67.7%	1.1%	60.8%	6.9%	77833	1.1%
General and Internal Medicine	22,410	65.1%	5.7%	56.7%	8.3%	98005	5.5%
Cancer	19,440	73.8%	0.4%	67.4%	6.3%	85627	0.7%
Surgery	16,961	71.9%	0.8%	49.1%	22.7%	32396	0.4%
Immunology	16,822	73.2%	0.5%	67.5%	5.7%	79388	0.3%
Cardiovascular System	15,011	68.7%	1.7%	50.4%	18.3%	41586	1.7%
All	145,536	71.5%|	1.5%	62.0%	9.5%	703565	1.3%
Mechanical Engineering	13,669	20.9%	2.9%	19.8%	1.1%	17620	0.5%
Computers	17,768	43.2%	1.9%	41.6%	1.6%	94350	0.7%
Electrical Engineering	30,271	40.4%	0.6%	35.7%	4.7%	65842	0.6%
Chemical Engineering	13,486	26.7%	1.0%	26.1%	0.6%	25857	0.4%
Materials Science	34,196	34.1%	1.8%	32.6%	1.4%	123535	0.9%
All	10,9390	34.8%	1.5%	32.5%	2.2%	327204	0.7%
Economics	12,300	41.0%	3.2%	40.2%	0.7%	63950	1.7%
General Social Science	26,28	40.3%	2.40%	39.6%	0.6%	11579	1.4%
Education	6,620	54.6%	6.8%	53.9%	0.%	49610	4.4%
LIS	2,330	62.1%	6.4%	59.5%	2.6%	20183	3.1%
All	23,878	46.7%	4.8%	45.8%	0.9%	145322	2.8%
Applied Physics	29,679	32.6%	1.2%	30.4%	2.1%	71050	0.6%
General Physics	36,595	29.0%	1.4%	27.7%	1.2%	94520	0.6%
Nuclear and Particle Physics	10,381	16.5%	0.8%	14.9%	1.5%	5225	0.8%
Optics	14,229	46.9%	0.8%	43.4%	3.50%	48614	0.5%
Solid State Physics	10,697	30.0%	1.6%	29.6%	0.4%	33385	0.9%
All	101,581	31.4%	1.2%	29.6%	1.7%	252794	0.6%
General Chemistry	23,144	29.9%	2.3%	28.8%	1.1%	70228	1.0%
Polymers	12,247	22.7%	3.2%	22.0%	0.7%	19478	1.6%
Physical Chemistry	36,329	35.4%	1.3%	31.0%	4.3%	85717	0.5%
Organic Chemistry	16,854	28.8%	1.1%	26.9%	1.8%	24190	0.6%
Analytical Chemistry	12,020	53.3%	1.3%	46.4%	6.9%	36767	0.9%
All	100,594	33.6%	1.6%	30.6%	3.0%	236380	0.8%
Total	480,979	45.6%	1.7%	41.1%	4.4%	1,665,265	1.1%

Appendix 7: Complete and merged categories for Mendeley readers' occupations.

Occupation provided by the Mendeley API	Merged	Individual readership
Assistant Professor	Assistant Professor	44,806
Lecturer		
Associate Professor	Associate Professor	27,283
Senior Lecturer		
Librarian	Librarian	7,476
Other Professions	Other Professions	44,183
PhD Student	PhD Student	567,122
Doctoral Student		
Postdoc	Postdoc	166,417
Professor	Professor	35,003
Researcher (at a non-Academic Institution)	Researcher (at a non-Academic Institution)	46,121
Researcher (at an Academic Institution)	Researcher (at an Academic Institution)	68,321
Student (Bachelor)	Student (Bachelor)	33,529
Student (Postgraduate)	Student (Postgraduate)	165,152
Student (Master)		

Appendix 8: Readers of WoS articles from Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry in Mendeley based on academic status for all papers regardless of % of available readership.

	%Engineering and Technology	%Social Science	%Physics	%Chemistry	%Clinical Medicine
PhD Students	55.4%	54.9%	51.7%	50.3%	39.1%
Postgrad Student	17.4%	19.5%	9.6%	11.1%	12.6%
Postdoc	8.5%	3.0%	18.2%	13.9%	17.4%
Researcher ac	3.8%	4.4%	5.5%	5.4%	6.9%
Assistant Professor	2.9%	4.9%	3.4%	3.1%	4.2%
Researcher non-ac	3.3%	1.7%	3.7%	5.8%	3.9%
Professor	2.2%	2.1%	2.9%	3.3%	3.3%
Bachelor Student	2.8%	3.0%	1.4%	2.4%	3.5%
Other Professions	1.7%	1.7%	1.1%	2.2%	5.9%
Associate Professor	1.8%	2.3%	2.2%	2.4%	2.4%
Librarian	0.2%	2.5%	0.2%	0.2%	0.75%
Total individual readership counts	244,097	97,191	192,222	177,909	457,954

Appendix 9: Readers of WoS articles from Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry in Mendeley based on academic status for with 66% reader counts.

	Engineering and Technology	Social Science	Physics	Chemistry	Clinical Medicine
PhD Students	56.3%	56.1%	52.1%	51.0%	39.5%
Postgrad Student	17.0%	18.3%	9.3%	10.7%	12.1%
Postdoc	8.3%	2.9%	18.2%	13.7%	17.2%
Researcher ac	3.9%	4.3%	5.6%	5.5%	7.0%
Researcher non-ac	3.3%	1.5%	3.7%	5.9%	4.2%
Assistant Professor	2.9%	4.9%	3.3%	3.0%	4.2%
Professor	2.2%	2.3%	3.0%	3.4%	3.6%
Bachelor Student	2.4%	2.9%	1.3%	2.2%	3.3%
Associate Professor	1.9%	2.4%	2.3%	2.5%	2.6%
Other Professions	1.5%	1.6%	1.1%	2.0%	5.6%
Librarian	0.2%	2.8%	0.2%	0.2%	0.7%
Total individual readership counts	194,128	60,874	159,507	142,919	302,814

Appendix 10: Readers of WoS articles from Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry in Mendeley based on academic status for with 100% reader counts.

	Engineering and Technology	Social Science	Physic s	Chemistry	Clinical Medicine
PhD Students	51.4%	46.9%	45.1%	45.2%	31.5%
Postgrad Student	15.5%	14.8%	9.6%	10.0%	13.8%
Postdoc	7.3%	3.5%	15.7%	11.8%	12.6%
Researcher ac	5.3%	5.6%	7.5%	7.2%	8.7%
Assistant Professor	4.5%	7.6%	5.0%	4.9%	6.1%
Researcher non-ac	4.2%	2.3%	4.7%	7.2%	5.4%
Professor	3.4%	4.6%	4.6%	4.8%	5.5%
Associate Professor	3.8%	5.2%	4.4%	4.1%	4.6%
Other Professions	2.0%	2.5%	1.6%	2.4%	7.2%
Bachelor Student	2.1%	3.1%	1.4%	2.0%	3.5%
Librarian	0.5%	3.7%	0.5%	0.4%	1.1%
Total individual readership counts	51,453	9,892	43,599	42,967	101,276

Appendix 11: Spearman correlations between WoS citations and Mendeley readership counts (non-zero only) for 2008 articles from Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry.

Main discipline	Sub-discipline	WoS citation median	Mendeley readership median	Correlation
Clinical Medicine	Neurology and Neurosurgery	10	7	.585**
	Pharmacology	9	4	.536**
	General and Internal Medicine	5	4	.563**
	Cancer	12	4	.604**
	Surgery	6	3	.451**
	Immunology	10	5	.573**
	Cardiovascular System	9	3	.592**
	All	9	4	.561**
Engineering and Technology	Mechanical Engineering	4	5	.533**
	Computers	3	7	.414**
	Electrical Engineering	4	4	.442**
	Chemical Engineering	7	5	.494**
	Materials Science	9	6	.682**
	All	5	5	.501**
Social Science	Economics	5	8	.629**
	General Social Science	3	8	.552**
	Education	4	9	.532**
	LIS	3	10	.546**
	All	4	8	.561**
Physics	Applied Physics	5	5	.566**
	General Physics	7	5	.595**
	Nuclear and Particle Physics	10	2	.325**
	Optics	6	5	.538**
	Solid State Physics	9	7	.628**
	All	7	5	.548**
Chemistry	General Chemistry	15	7	.648**
	Polymers	10	5	.595**
	Physical Chemistry	10	5	.527**
	Organic Chemistry	10	4	.423**
	Analytical Chemistry	10	4	.528**
	All	11	5	.554**

Appendix 12: Spearman correlations between Mendeley readership counts and citations based on occupation for Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry articles for all articles regardless of percentage of readership availability.

		Clinical Medicine	Physics	Engineering and Technology	Chemistry	Social Sciences
Professor	Spearman's rho	.404**	.439**	.482**	.435**	.485**
	N	9,549	3,345	3,142	3,550	1,048
Associate Professor	Spearman's rho	.292**	.337**	.288**	.379**	.345**
	N	8,358	3,012	3,190	3,018	1,328
Assistant Professor	Spearman's rho	.406**	.427**	.381**	.420**	.471**
	N	11,931	3,930	4,353	3,587	2,284
Researcher (at an Academic Institution)	Spearman's rho	.414**	.403**	.358**	.371**	.478**
	N	17,702	6,161	5,533	5,813	1,829
Researcher (at a non-Academic Institution)	Spearman's rho	.418**	.411**	.410**	.368**	.552**
	N	9,908	3,727	4,159	5,273	725
Post Doc	Spearman's rho	.446**	.501**	.518**	.464**	.493**
	N	30,274	14,014	9,210	11,626	1,420
Ph.D. Student	Spearman's rho	.435**	.518**	.458**	.485**	.523**
	N	53,169	23,197	29,064	23,859	8,990
Student (Postgraduate)	Spearman's rho	.326**	.417**	.328**	.392**	.455**
	N	31,106	9,723	17,141	10,715	5,765
Student (Bachelor)	Spearman's rho	.245**	.217**	.261**	.264**	.354**
	N	10,990	2,000	4,045	3,147	1,510
Other Professional	Spearman's rho	.315**	.219**	.122**	.171**	.294**
	N	16,861	1,734	2,937	2,700	1,042
Librarian	Spearman's rho	.078**	-0.003	-0.05	-0.033	.229**
	N	2,808	415	480	370	768

**Significant at p = 0.01.

Appendix 13: Spearman correlations between Mendeley readership counts and citations based on occupation for Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry articles for articles with at least 66% readership availability.

Occupation		Clinical Medicine	Physics	Engineering and Technology	Chemistry	Social Sciences 66%
Professor	Spearman's rho	.323**	.403**	.433**	.378**	.435**
	N	7,778	3,053	2,764	3,142	826
Associate Professor	Spearman's rho	.228**	.310**	.270**	.335**	.321**
	N	6,399	2,720	2,832	2,647	977
Assistant Professor	Spearman's rho	.347**	.391**	.361**	.379**	.461**
	N	8,724	3,414	3,699	3,012	1,607
Researcher (at an Academic Institution)	Spearman's rho	.328**	.371**	.329**	.324**	.419**
	N	13,660	5,516	4,829	5,074	1,333
Researcher (at a non-Academic Institution)	Spearman's rho	.334**	.357**	.357**	.315**	.461**
	N	7,828	3,276	3,626	4,643	532
Post Doc	Spearman's rho	.397**	.475**	.487**	.438**	.429**
	N	22,413	12,140	7,796	9,673	1,005
Ph.D. Student	Spearman's rho	.405**	.504**	.446**	.467**	.499**
	N	39,887	20,368	25,314	20,361	6,531
Student (Postgraduate)	Spearman's rho	.262**	.388**	.310**	.358**	.396**
	N	22,234	8,234	14,548	8,759	3,977
Student (Bachelor)	Spearman's rho	.173**	.195**	.232**	.237**	.298**
	N	7,489	1,603	3,263	2,451	1,053
Other Professional	Spearman's rho	.225**	.185**	.101**	.137**	.203**
	N	11,683	1,397	2,313	2,068	698
Librarian	Spearman's rho	0.032	0.002	-0.060	-0.054	.176**
	N	1,886	356	386	309	567

**Significant at p = 0.01.

Appendix 14: Spearman correlations between Mendeley readership counts and citations based on occupation for Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry articles for articles with 100% readership availability.

Occupation		Clinical Medicine	Physics	Engineering and Technology	Chemistry	Social Sciences
Professor	Spearman's rho	.086**	.094**	.149**	.095**	.130**
	N	5,059	1,764	1,537	1,778	394
Associate Professor	Spearman's rho	.052**	0.024	0.012	.066**	-0.008
	N	4,289	1,729	1,798	1,637	467
Assistant Professor	Spearman's rho	.096**	.059**	.044*	.091**	.136**
	N	5,578	1,954	2,102	1,905	650
Researcher (at an Academic Institution)	Spearman's rho	.097**	.104**	.108**	.052**	0.029
	N	7,473	2,740	2,401	2,662	464
Researcher (at a non-Academic Institution)	Spearman's rho	.109**	.058*	.060**	.085**	0.069
	N	4,578	1,664	1,847	2,376	202
Post Doc	Spearman's rho	.145**	.143**	.144**	.161**	.127*
	N	10,009	5,022	3,181	3,951	316
Ph.D. Student	Spearman's rho	.212**	.233**	.226**	.225**	.201**
	N	19,985	9,831	12,591	9,710	2,307
Student (Postgraduate)	Spearman's rho	.077**	.061**	.095**	.115**	.107**
	N	11,267	3,497	6,061	3,473	1,075
Student (Bachelor)	Spearman's rho	0.001	-0.056	0.047	0.063	.153*
	N	3,065	547	974	799	276
Other Professional	Spearman's rho	.082**	-0.016	0.019	-0.003	-0.079
	N	6,036	660	968	923	236
Librarian	Spearman's rho	0.004	-0.005	-0.036	-0.097	-0.079
	N	1,046	211	233	179	238

**Significant at p = 0.01.

Appendix 15: Descriptive statistics for Mendeley users who published personal webpage in their Mendeley profiles

Field	Discipline	Number and % of users in Mendeley directory	Number of users who have contact information in Mendeley profile*	Extracted emails
Art and Humanities	Arts and Literature	4279 (2.3%)	554	169 (2.8%)
	Humanities	4575 (2.4%)	397	94 (1.5%)
	Law	1662 (0.9%)	198	91 (1.5%)
	Linguistics	2174 (1.2%)	299	87(1.4%)
	Design	2480 (1.3%)	427	92 (1.5%)
	Philosophy	1690 (0.9%)	175	65(1.1%)
	All	*14380 (7.6%)*	*2050*	*598 (9.8%)*
Basic science	Environmental sciences	6152 (3.3%)	704	238(3.9%)
	Chemistry	6030 (3.2%)	402	163(2.7%)
	Computer and information science	27491 (14.6%)	3500	842 (13.8%)
	Earth science	4445 (2.4)	587	176 (2.9%)
	Materials science	2631 (1.4%)	277	67 (1.1%)
	Mathematics	2442 (1.3%)	334	117(1.9%)
	Physics and Astronomy	8090 (4.3%)	907	307(5.0%)
	All	*43727 (23.2%)*	*6711*	*1910 (31.2%)*
Engineering	Electrical and electronic engineering	5842 (3.1%)	674	206 (3.4%)
	Engineering	13387 (7.1%)	1273	523 (8.5%)
	All	*19229 (10.2%)*	*1947*	*729 (11.9%)*
Medicine and biology	Biological Sciences	31216 (16.6%)	3323	966 (15.8%)
	Medicine	17665 (9.4%)	1117	525(8.6%)
	All	*48881 (26.0%)*	*4440*	1491(24.4%)
Social Sciences	Business Administration	8552 (4.5%)	583	224 (3.7%)
	Economics	4101 (2.2%)	491	154 (2.5%)
	Education	10047 (5.3%)	280	74 (1.2%)
	Management Science	3428 (1.8%)	386	89 (1.5%)
	Psychology	8981 (4.8%)	1243	366 (6.0%)
	social-sciences	13398 (7.1%)	1765	470 (7.7%)
	Sports and Recreation	1316 (0.7%)	63	17(0.3%)
	All	*49823 (26.5%)*	*4811*	1394 (22.8%)
Total	*Total*	*188100*	*19959*	*6122*

Appendix 16: The full version of the questionnaire.

Please complete this short survey into reasons for using Mendeley. All responses to the questions are voluntary, anonymous and confidential. For further information about this study, please contact me at e.mohammadi@wlv.ac.uk

1. What is your current position?

Full professor

Associate professor / reader / senior lecturer

Assistant professor / lecturer

Researcher (e.g., postdoctoral, or temporary researcher)

PhD student

Masters student

Undergraduate student

Other (please specify)

2. What is your main broad discipline?

Medical Science

Basic Science (e.g., physics, chemistry, maths)

Engineering

Social Science

Arts and Humanities

Other (please specify)

3. How do you use Mendeley? (Please select all those that apply)

As a reference manager or to maintain a bibliography for academic activities (e.g., research, publication or teaching)

To publicise your own publications

As a social networking site, communicating with other users

As a database to search for publications listed within Mendeley

4. <u>Have you added any documents to your Mendeley personal library? (Excluding your own publications)?</u>

Yes

No

5. <u>What are the main purposes for adding records (e.g., articles) to your Mendeley personal library? (Please select all those that apply)</u>

To cite them in my publications (e.g., papers, books)

To cite in my thesis / dissertation

To use them for teaching purposes (e.g., syllabi, reading lists and bibliographies for students)

To use them in my assignments for a course that I am taking

For professional use (e.g., to keep track of research relevant to my medical or engineering job)

Other (please specify)

6. <u>Approximately what proportion of the items in your Mendeley personal library have you read OR do you intend to eventually read? Please estimate as well as you can.</u>

All

At least half but not all

At least one but less than half.

None

Appendix 17: Ethical approval for reasons for Using Mendeley survey.

Personal information

Name of Student: Ehsan Mohammadi

Email address: e.mohammadi@wlv.ac.uk

Degree Programme: PhD

Director of studies: Professor Mike Thelwall

Research questions

This survey aims to discover motivations for using Mendeley.com, a reference manager website. The following research question drives this investigation.

1. How and why do people use Mendeley?

Participants and Data collection

A sample of Mendeley users who published their contact information will be recruited with the following procedure;

1. Mendeley profiles will be extracted based on academic disciplines from the Mendeley people directory http://www.mendeley.com/directory (around 200,000 profiles).
2. In order to differentiate users who published their personal webpages in Mendeley profile publicly, using Webometric Analyst (lexiurl.wlv.ac.uk) the data was collected from Bing (http://www.bing.com/) API (Application programming interface). The following query was used

"Webpage:" "two last keywords of the Mendeley Profile URL" site: "http://www.mendeley.com/profiles/"

Example: "Webpage:" "kayvan-kousha" site: http://www.mendeley.com/profiles/

3. Contact information (emails) of a sample of the users, around 10,000, will be extracted manually from their personal webpages.

4. The chosen sample will be contacted. A copy of the standard message to be sent to Mendeley users is below and a copy of the questionnaire is also attached. Users will be given a link to complete the questionnaire in SurveyMonkey, which will be anonymous.

Data analysis

No information will be collected about the identity of those who complete the questionnaire. The responses will be stored and hence analysed anonymously for a simple descriptive analysis of the common reasons for using Mendeley.

Standard Message for participants

Subject: Survey of reasons for using Mendeley

Dear [name to be completed],

Please could you complete a brief survey (5 questions) about why you use Mendeley? This will take about 2 minutes of your time.

https://www.surveymonkey.com/s/S86L2XX

Your answers will be fully anonymous and will be used to help understand whether Mendeley can be used for research evaluation. Please click the link above to take the survey.

This survey has received ethical approval from the University of Wolverhampton Research Institute for Information and Language Processing.

Thank you very much in advance,

Ehsan Mohammadi

PhD student

University of Wolverhampton, UK

Appendix 18: Respondents to the survey based at the level of sub-disciplines.

Field	Discipline	Number and % of users in Mendeley directory	Number of recruited users (bounced are excluded)	% of recruited users in the sample	Number of respondents	Response rate
Arts and Humanities	Arts and Literature	4,279 (2.3%)	162	2.7%	17	10.5%
	Humanities	4,575 (2.4%)	92	1.6%	7	7.6%
	Law	1,662 (0.9%)	90	1.5%	22	24.4%
	Linguistics	2,174 (1.2%)	84	1.4%	15	17.9%
	Design	2,480 (1.3%)	90	1.5%	6	6.7%
	Philosophy	1,690 (0.9%)	64	1.1%	9	14.1%
	All	14,380 (7.6%)	582	9.8%	76	13.1%
Basic Science	Environmental Sciences	6,152 (3.3%)	228	3.8%	23	10.1%
	Chemistry	6,030 (3.2%)	159	2.7%	29	18.2%
	Computer and Information Science	27,491 (14.6%)	814	13.7%	97	11.9%
	Earth Science	4445 (2.4)	172	2.9%	32	18.6%
	Materials Science	2631 (1.4%)	66	1.1%	14	21.2%
	Mathematics	2,442 (1.3%)	113	1.9%	20	17.7%
	Physics and Astronomy	8,090 (4.3%)	291	4.9%	38	13.1%
	All	43,727 (23.2%)	1,843	31.1%	253	13.7%
Engineering	Electrical and Electronic Engineering	5,842 (3.1%)	199	3.4%	33	16.6%
	Engineering	13,387 (7.1%)	512	8.6%	89	17.4%
	All	19,229 (10.2%)	711	12.0%	122	17.2%
Medicine and Biology	Biological Sciences	31,216 (16.6%)	938	15.8%	118	12.6%
	Medicine	17,665 (9.4%)	502	8.5%	68	13.5%
	All	48,881 (26.0%)	1,440	24.3%	186	12.9%
Social Sciences	Business Administration	8,552 (4.5%)	218	3.7%	38	17.4%
	Economics	4,101 (2.2%)	148	2.5%	23	15.5%
	Education	10047 (5.3%)	72	1.2%	19	26.4%
	Management Science	3,428 (1.8%)	84	1.4%	14	16.7%
	Psychology	8,981 (4.8%)	356	6.0%	67	18.8%
	Social-Sciences	13,398 (7.1%)	457	7.7%	63	13.8%
	Sports and Recreation	1,316 (0.7%)	16	0.3%	3	18.8%
	All	49,823 (26.5%)	1,351	22.8%	227	16.8%
Total		*188,100*	*5,927*	*100.0%*	*864*	*14.6%*

Appendix 19: Chi-Square tests of purposes for using Mendeley across different disciplines.

Table 6.4. A chi-square test of motivations for using Mendeley across different disciplines.

Reason for using Mendeley/ discipline	Arts and Humanities	Basic Science	Engineering	Medical Science	Social Science	p value
As a reference manager	48	202	198	104	200	0.691
To publicize your own publications	11	58	57	16	64	
As a social networking site	9	28	32	19	38	
As a database to search for publications	18	63	61	34	71	

Table 6.5. A chi-square test for using Mendeley as a reference manager tool across different disciplines.

Discipline	Yes	No	p value
Arts and Humanities	48	38	0.102
Basic Science	202	149	
Engineering	198	150	
Medical Science	104	69	
Social Science	200	102	

Table 6.6. A chi-square test for using Mendeley to publicize user's publications across different disciplines.

Discipline	Yes	No	p value
Arts and Humanities	11	75	0.014
Basic Science	58	293	
Engineering	57	291	
Medical Science	16	157	
Social Science	64	238	

Table 6.7. A chi-square test for using Mendeley as a social networking site across different disciplines**.**

Discipline	Yes	No	p value
Arts and Humanities	9	77	0.742
Basic Science	28	223	
Engineering	32	316	
Medical Science	19	154	
Social Science	38	264	

Table 6.8. A chi-square test for using Mendeley as a database to search for publications across different disciplines.

Discipline	Yes	No	p value
Arts and Humanities	18	68	0.333
Basic Science	63	288	
Engineering	61	287	
Medical Science	34	139	
Social Science	71	231	

Appendix 20: Chi square tests of purposes for using Mendeley for different type of users.

Table 6.9. A chi-Square test of motivations for using Mendeley by user occupation.

Reason of using Mendeley/ Occupation	Assistant professor	Associate professor	professor	Other Professions	PhD student	Researcher	Undergraduate and masters students	P-Value
As a reference manager	102 (57%)	92 (50%)	76 (48%)	43 (51%)	217 (64%)	195 (59%)	27 (49%)	0.025
To publicize your own publications	31 (17%)	37 (20%)	28 (18%)	12 (14%)	37 (11%)	57 (17%)	5 (9%)	
As a social networking site	16 (9%)	20 (11%)	18 (11%)	12 (14%)	32 (9%)	22 (7%)	7 (13%)	
As a database to search for publications	31 (17%)	34 (19%)	36 (23%)	18 (21%)	54 (16%)	58 (17%)	16 (29%)	
Total	180	183	158	85	340	332	55	

Table 6.10. A chi-square test for using Mendeley as a reference manager for different user occupations.

Occupation	Yes	No	p value
Assistant professor	102	78	0.006
Associate professor	92	91	
Professor	76	82	
Other Professions	43	42	
PhD student	217	123	
Researcher	195	137	
Undergraduate and masters students	27	28	

Table 6.11. A chi-square for using Mendeley to publicize a user's publications by user occupation.

Occupation	Yes	No	p value
Assistant professor	31	149	0.057
Associate professor	37	146	
Professor	28	130	
Other Professions	12	73	
PhD student	37	303	
Researcher	57	275	
Undergraduate and masters students	5	50	

Table 6.12. A chi-square test for using Mendeley as a social networking site for different user occupations.

Occupation	Yes	No	p value
Assistant professor	16	164	0.310
Associate professor	20	163	
Professor	18	140	
Other Professions	12	73	
PhD student	32	308	
Researcher	22	310	
Undergraduate and masters students	7	48	

Table 6.13. A chi-square for using Mendeley as a database to search for publications for different user occupations.

Occupation	Yes	No	p value
Assistant professor	31	149	0.273
Associate professor	34	149	
Professor	36	122	
Other Professions	18	67	
PhD student	54	286	
Researcher	58	274	
Undergraduate and masters students	16	39	

Appendix 21: A chi-Square test of motivations for bookmaking documents in Mendeley across different disciplines.

Table 6.14. A chi-Square test of all motivations for bookmaking documents in Mendeley across different disciplines.

Motivations of bookmarking / discipline	Arts and Humanities	Basic Science	Engineering	Medical Science	Social Science	p value
To cite them in my publications (e.g., papers, books)	32 (33%)	153 (42%)	150 (38%)	74 (39%)	165 (40%)	0.001
To cite in my thesis / dissertation	23 (24%)	74 (20%)	89 (23%)	21 (11%)	73 (18%)	
To use them for teaching purposes	12 (13%)	33 (9%)	39 (10%)	26 (14%)	63 (15%)	
To use them in my assignments for a course that I am taking	9 (9%)	14 (4%)	22 (6%)	9 (5%)	33 (8%)	
For professional use (job)	20 (21%)	90 (25%)	94 (24%)	60 (32%)	78 (19%)	
Total	96	364	394	190	412	

Table 6.15. A chi-square test for citing bookmarked documents in future publications (e.g., papers, books) across different disciplines.

Disciplines	Yes	No	p value
Arts and Humanities	32	10	0.023
Basic Science	153	32	
Engineering	150	25	
Medical Science	74	21	
Social Science	165	17	

Table 6.16. A chi-square test for citing bookmarked documents in theses or dissertations across different disciplines.

Disciplines	Yes	No	p value
Arts and Humanities	23	19	0.000
Basic Science	74	111	
Engineering	89	86	
Medical Science	21	74	
Social Science	73	109	

Table 6.17. A chi-square test for using bookmarked documents in teaching activities across different disciplines.

Discipline	Yes	No	p value
Arts and Humanities	12	30	0.004
Basic Science	33	152	
Engineering	39	136	
Medical Science	26	69	
Social Science	63	119	

Table 6.18. A chi-square test for using bookmarked documents in assignments across different disciplines.

Discipline	Yes	No	p value
Arts and Humanities	9	33	0.011
Basic Science	14	171	
Engineering	22	153	
Medical Science	9	86	
Social Science	33	149	

Table 6.19. A chi-square test o using bookmarked documents in professional (job) activities across different disciplines.

Discipline	Yes	No	p value
Arts and Humanities	20	22	0.021
Basic Science	90	95	
Engineering	94	81	
Medical Science	60	35	
Social Science	78	104	

Appendix 22: The chi-Square tests for motivations of bookmaking documents in Mendeley for different type of users.

Table 6.20. A chi-Square test of all motivations for bookmaking documents in Mendeley for different user occupations.

Motivations of bookmarking / Occupation	Assistant professor	Associate professor	professor	Other Professions	PhD student	Researcher	Undergraduate and masters students	p value
Future citation (publications and thesis)	90 (51%)	74 (48%)	57 (40%)	20 (31%)	353 (69%)	203 (6%)	26 (65%)	0
Educational and teaching activities (assignment and teachings)	38 (22%)	37 (24%)	44 (31%)	15 (23%)	81 (16%)	38 (11%)	9 (23%)	
For professional use (job)	48 (27%)	44 (28%)	42 (29%)	29 (45%)	77 (15%)	98 (29%)	5 (13%)	
Total	176	155	143	64	511	339	40	

Table 6.21. A chi-square test for using bookmarked documents for future citation in papers, books and thesis and dissertation for different user occupations.

Occupation	Yes	NO	p value
Assistant professor	90	86	0.000
Associate professor	74	81	
Professor	57	86	
Other Professions	20	44	
PhD student	353	158	
Researcher	203	136	
Undergraduate and masters students	26	14	

Table 6.22. A chi-square test for using bookmarked articles in educational and teaching activities for different user occupations.

Profession	Yes	NO	p value
Assistant professor	38	138	0.000
Associate professor	37	118	
Professor	44	99	
Other Professions	15	49	
PhD student	81	430	
Researcher	38	301	
Undergraduate and masters students	9	31	

Table 6.23. The results of the Chi-Square test (p Value) for using bookmarked in professional (job) activities for different user occupations.

Occupation	Yes	NO	p value
Assistant professor	48	128	0.000
Associate professor	44	111	
Professor	42	101	
Other Professions	29	35	
PhD student	77	434	
Researcher	98	241	
Undergraduate and masters students	5	35	

Appendix 23: A chi-Square test for the proportion of the items read from Mendeley personal libraries across different disciplines.

Proportion of the items have read from Mendeley personal libraries	%Social Science	%Arts and Humanities	%Basic Science	%Engineering	%Medical Science	P value
All	24%	43%	24%	31%	26%	0.282
At least half but not all	59%	43%	55%	51%	57%	
At least one but less than half.	17%	14%	21%	17%	16%	
None	0%	0%	0%	1%	1%	
Total	185	42	182	173	97	

Appendix 24: A cross tabulation of Engineering and Technology papers based on their Mendeley readership and WoS citations.

WoS Citations / Mendeley Readership	% of papers with zero citations	% of papers with 1-5 citations	% of papers with 6--10 citations	% of papers with 11--20 citations	% of papers with 21 and more citations
% of papers with 1—5 readership counts	9.1%	27.7%	9.4%	4.5%	1.3%
% of papers with 6--10 readership counts	1.4%	8.9%	6.1%	4.9%	2.2%
% of papers with 11 and over readership counts	0.8%	4.9%	4.7%	5.9%	8.2%
Total	11.4%	41.4%	20.2%	15.4%	11.6%

Appendix 25: A cross tabulation of Social Science papers based on their Mendeley readership and WoS citations.

WoS Citations / Mendeley Readership	% of papers with zero citations	% of papers with 1-5 citations	% of papers with 6--10 citations	% of papers with 11--20 citations	% of papers with 21 and more citations
% of papers with 1—5 readership counts	10.0%	20.5%	3.6%	1.3%	0.3%
% of papers with 6--10 readership counts	2.9%	13.2%	4.6%	1.9%	0.5%
% of papers with 11 and over readership counts	1.6%	13.7%	11.2%	9.2%	5.7%
Total	14.5%	47.4%	19.3%	12.4%	6.5%

Appendix 26: A cross tabulation of Clinical Medicine papers based on their Mendeley readership and WoS citations.

WoS Citations / Mendeley Readership	% of papers with zero citations	% of papers with 1-5 citations	% of papers with 6--10 citations	% of papers with 11--20 citations	% of papers with 21 and more citations
% of papers with 1—5 readership counts	4.4%	24.7%	14.9%	11.5%	4.2%
% of papers with 6--10 readership counts	0.3%	3.8%	4.8%	6.6%	5.5%
% of papers with 11 and over readership counts	0.1%	1.4%	2.2%	4.6%	10.9%
Total	4.9%	29.9%	21.9%	22.7%	20.6%

Appendix 27: A cross tabulation of Chemistry papers based on their Mendeley readership and WoS citations.

WoS Citations / Mendeley Readership	% of papers with zero citations	% of papers with 1-5 citations	% of papers with 6--10 citations	% of papers with 11--20 citations	% of papers with 21 and more citations
% of papers with 1—5 readership counts	3.2%	18.7%	13.8%	11.9%	5.2%
% of papers with 6--10 readership counts	0.2%	3.9%	5.9%	8.5%	6.9%
% of papers with 11 and over readership counts	0.1%	0.9%	2.2%	5.7%	13.0%
Total	3.4%	23.5%	21.8%	26.1%	25.2%

Appendix 28: A cross tabulation of Physics papers based on their Mendeley readership and WoS citations.

WoS Citations / Mendeley Readership	% of papers with zero citations	% of papers with 1-5 citations	% of papers with 6--10 citations	% of papers with 11--20 citations	% of papers with 21 and more citations
% of papers with 1—5 readership counts	7.0%	26.7%	11.0%	6.3%	2.6%
% of papers with 6--10 readership counts	1.0%	7.1%	6.4%	5.6%	2.3%
% of papers with 11 and over readership counts	0.3%	2.3%	4.0%	7.2%	10.2%
Total	8.3%	36.0%	21.4%	19.1%	15.2%

Appendix 29: Top 5 papers in terms of Mendeley readers but zero citations for Clinical Medicine, Engineering and Technology, Social Science, Physics and Chemistry papers.

Article	**Citations**	**Mendeley readers**	**Discipline**
Cormode, G, 2008, How NOT to review a paper The tools and techniques of the adversarial reviewer, SIGMOD RECORD	0	108	Engineering
Adam, JA, 2008, Geometric optics and rainbows: generalization of a result by Huygens, APPLIED OPTICS	0	106	Physics
Sachs, J, 2008, The end of poverty: economic possibilities for our time, EUROPEAN JOURNAL OF DENTAL EDUCATION	0	87	Social science
Williams, S, 2008, HOW WE FOUND THE MISSING MEMRISTOR, IEEE SPECTRUM	0	79	Engineering
Nightingale, EB., Veeraraghavan, K., Chen, PM., Flinn, J., 2008, Rethink the sync, ACM TRANSACTIONS ON COMPUTER SYSTEMS	0	76	Engineering
Fettke, P, 2008, Business Process Modeling Notation, WIRTSCHAFTSINFORMATIK	0	70	Engineering
Preda, MD., Christodorescu, M., Jha, S., Debray, S., 2008, A semantics-based approach to malware detection, ACM TRANSACTIONS ON PROGRAMMING LANGUAGES AND SYSTEMS	0	56	Engineering
Bayer, P., Ross, SL., Topa, G., 2008, Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes, JOURNAL OF POLITICAL ECONOMY	0	51	Social science
McGrath, M., 2008, Interlending and document supply: a review of the recent literature: 62, INTERLENDING & DOCUMENT SUPPLY	0	39	Social science
Altevogt, BM., Hanson, SL., Leshner, A., 2008,Molecules to Minds: Grand Challenges for the 21st Century, NEURON	0	36	Clinical Medicine
Fitzpatrick, DA., Logue, ME., Butler, G., 2008, Evidence of recent interkingdom horizontal gene transfer between bacteria and Candida parapsilosis, BMC IMMUNOLOGY	0	34	Clinical Medicine
Bhushan, B., Jung, YC., 2008, Wetting, adhesion and friction of superhydrophobic and hydrophilic leaves and fabricated micro/nanopatterned surfaces, JOURNAL OF PHYSICS-CONDENSED MATTER	0	33	Physics
Chen, CC., Wu, J., Yang, SC., Tsou, HY., 2008, Importance of diversified leadership roles in improving team effectiveness in a virtual collaboration learning environment, EDUCATIONAL TECHNOLOGY & SOCIETY	0	33	Social science
Duchesne, S., Jannin, P., 2008, Proposing a manuscript peer-review checklist, NEUROIMAGE	0	32	Clinical Medicine
Nicolasora, N., Kaul, DR., 2008, Infectious disease emergencies, MEDICAL CLINICS OF NORTH AMERICA	0	32	Clinical Medicine
MacGowan, AP., 2008, Clinical implications of antimicrobial resistance for therapy, JOURNAL OF ANTIMICROBIAL CHEMOTHERAPY	0	32	Clinical Medicine
Benka, SG, 2008, Who is listening? What do they hear?, PHYSICS TODAY	0	31	Physics

Article	Citations	Mendeley readers	Discipline
Furtak, EM., Ruiz-Primo, MA., 2008, Making students' thinking explicit in writing and discussion: An analysis of formative assessment prompts, SCIENCE EDUCATION	0	31	Social science
Nosonovsky, M., Bhushan, B., 2008, Roughness-induced superhydrophobicity: a way to design non-adhesive surfaces, JOURNAL OF PHYSICS-CONDENSED MATTER	0	27	Physics
Abell, J., Moustakas, TD., 2008, The role of dislocations as nonradiative recombination centers in InGaN quantum wells, APPLIED PHYSICS LETTERS	0	23	Physics
Riess, I., 2008, Solid State Electrochemistry, ISRAEL JOURNAL OF CHEMISTRY	0	17	Chemistry
Han, J., Chen, JM., Zhou, XW., Lin, Y., Zhang, JB., Jia, JG., 2008, Dye-sensitized solid-state solar cells fabricated by screen-printed TiO2 thin film with addition of polystyrene balls,Dye-sensitized solid-state solar cells fabricated by screen-printed TiO2 thin film with addition of polystyrene balls	0	17	Chemistry
Mullin, JB., 2008, Innovation in crystal growth: A personal perspective, JOURNAL OF CRYSTAL GROWTH	0	13	Chemistry
Lopes, MC., de Oliveira, CP., Pereira, EC., 2008, Computational modeling of the template-assisted deposition of nanowires, ELECTROCHIMICA ACTA	0	12	Chemistry
Dagbert, C., Meylheuc, T., Bellon-Fontaine, MN., 2008, Pit formation on stainless steel surfaces pre-treated with biosurfactants produced by Pseudomonas fluorescens, ELECTROCHIMICA ACTA	0	12	Chemistry